GENE INTERACTIONS WITH AGOUTI SIGNALING PROTEIN PRODUCE COMPLEX PIGMENTATION PHENOTYPES IN THE DOMESTIC DOG

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Animal and Poultry Science
University of Saskatchewan
Saskatoon, SK
Canada

Dayna Lee Dreger, M.Sc.

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work, or in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying, publication, or use of this thesis or parts of it thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make use of material in this thesis in whole or in part should be addressed to:

Head of the Department of Animal and Poultry Science
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5A8

ABSTRACT

Much of what is understood about canine coat color genetics focuses on single gene inheritance. However, within the traditional color patterns, there is noticeable variation, suggesting modification of known coat color alleles. Alleles of the *Agouti Signaling Protein* (*ASIP*) gene control the temporal and spatial expression of red-based phaeomelanin and black-based eumelanin pigments. While up to six *ASIP* alleles have been predicted by previous studies, only two alleles, a^y fawn and *a* recessive black, have been identified as polymorphisms within *ASIP*. Through sequence analysis, a SINE insertion in intron 1 of *ASIP* was detected that segregates with the black-and-tan, saddle tan and recessive black phenotypes. Together with the diagnostic tests for the a^y and *a* alleles, detection of the SINE allows for complete genotyping of the four common *ASIP* alleles: a^y, a^w, a^t, *a*.

Salukis and Afghan Hounds exhibit a phenotype known as grizzle or domino. This phenotype is similar to the black-and-tan phenotype, also present in both breeds, though the phaeomelanin points extend further up the limbs, onto the ventral surface, and form a widow's peak on the face. Sequencing of the *Melanocortin 1 Receptor* (*MC1R*) gene identified a G78V polymorphism that segregates with the grizzle/domino phenotype. Further genotype analysis showed that expression of the grizzle/domino pattern requires a genotype interaction involving *ASIP*, *MC1R*, and *Canine Beta-Defensin 103* (*DEFB103*).

The saddle tan phenotype is present in scent hounds, terriers, and herding breeds, and is characterized by phaeomelanin pigment on the limbs, often extending to the midline of the back, and encompassing the entire head. Genome wide association study analysis, fine mapping, and sequence analysis identified a 16 bp tandem duplication in intron 5 of the *hnRNP-Associated with Lethal Yellow* (*RALY*) gene that segregates with the black-and-tan phenotype, versus saddle tan, in Basset Hounds and Pembroke Welsh Corgis. In breeds that never have the saddle tan phenotype, but frequently have the black-and-tan phenotype, the *RALY* duplication does not segregate with black-and-tan. This, together with further genotype analysis, suggests a gene interaction of *ASIP*, *MC1R*, *DEFB103*, *RALY*, and an additional modifier gene is required for expression of saddle tan or black-and-tan.

ACKNOWLEDGEMENTS

I'd like to extend thanks to my graduate committee, Dr. Fiona Buchanan, Dr. Yves Plante, Dr. Bill Roesler, and chairs Dr. Bernard Laarveld and Dr. Tim Mutzvangwa, for the guidance, support, critiques, and encouragement every step of the way.

Above all, I wish to thank my supervisor, Dr. Sheila Schmutz. I've had the honor of learning from you for 7 years now, and in that time you have taught me to believe in myself, to speak up for what I know even when I might be terrified to do so, that it's alright to brag now and then, to think outside the box, and that I'm always capable of more than I think I am.

I'd like to also take this opportunity to thank my parents for always acting as if every little accomplishment is brag-worthy. Thanks also have to go to my friends, in particular Natasha M., Judy M., Sheila G., Lisa M., and Kaley P., for the distractions, the laughs, the nights out, the confidence boosts, and for always keeping me humble and reminding me that "it's just science!"

Funding for my research was provided by a NSERC Discovery grant. The saddle tan research was done in collaboration with Dr. Heidi Parker and Dr. Elaine Ostrander of the National Institutes of Health in Bethesda, Maryland. Thanks to the Basset Hound Club of America 2010 National Specialty organizing committee for allowing me to attend their event in order to collect Basset Hound samples for the saddle tan research. Thanks to Casey Gonda and Margot Keast for arranging the Saluki samples required for the grizzle/domino portion of my research, and for providing insight into the color patterns present in the Saluki breed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

α-MSH	Alpha Melanocyte Stimulating Hormone
°C	Degrees Celsius
µl	Microlitre
ASIP	*Agouti Signaling Protein*
ATOX1	*Antioxidant Protein 1*
ATRN	*Attractin*
BAC	Bacterial Artificial Chromosome
bp	Base Pair
BPIFB2	*BPI fold-containing family B*
DEFB103	*Canine Beta Defensin 103*
CBFA2T2	*Core-Binding Factor*
CDK5	*Cyclin-Dependent Kinase 5*
cDNA	Complementary deoxyribonucleic acid
CFA	*Canis familiaris* autosome
ChE	Human Cholinesterase
cM	Centimorgan
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
E2F1	*E2F transcription factor 1*
EIF2B	*Eukaryotic Initiation Factor 2B*
FCI	Fédération Cynologique Internationale
FGF5	*Fibroblast Growth Factor 5*
GWAS	Genome Wide Association Study
hnRNP	Heterogenous nuclear ribonucleoprotein
kb	Kilobase
KRT71	*Keratin 71*
LD	Linkage disequilibrium
LINE	Long Interspersed Nuclear Element
LOD	Log of the odds
Mb	Megabase

MC1R	*Melanocortin 1 Receptor*
$MgCl_2$	Magnesium chloride
MITF	*Micro-ophthalmia Associated Transcription Factor*
min	Minutes
mL	Millilitre
mM	Millimolar
mRNA	Messenger ribonucleic acid
PCR	Polymerase Chain Reaction
PCR-RFLP	Polymerase Chain Reaction Restriction Fragment Length Polymorphism
PLUNC	*Palate Lung and Nasal Epithelium Clone Protein*
pmol	Picomole
PNPLA1	*Patatin-like Phospholipase Domain-containing Protein 1*
RALY	*hnRNP Associated with Lethal Yellow*
RNA	Ribonucleic acid
RPS20	*Ribosome Protein S20*
RSPO2	*R-spondin 2*
s	Seconds
SILV	also *PMEL17*, *Melanocyte Protein 17*
SINE	Short Interspersed Nuclear Element
SNP	Single nucleotide polymorphism
SNTA	*α-Syntrophin*
TBX15	*T-box 15*
TMHMM	Trans-membrane Hidden Markov Model
UTR	Untranslated Region
VNTR	Variable Number Tandem Repeat

1.0 GENERAL INTRODUCTION

Observing the inheritance of pigmentation patterns is a valuable resource for genetic research. This highly visual trait allows for easy identification of phenotype, and promotes collaboration of researchers with domestic animal breeders, owners, and veterinarians. Additionally, coat color genetics provides an ideal teaching tool for demonstrating the fundamentals of genetic inheritance to students of all levels, by utilizing readily identifiable, yet highly variable, external traits of a familiar domestic species, such as the dog.

Domestication and selective breeding of dogs has created a species that exhibits a vast range of phenotypic variation, both within and between very specialized and closely monitored pure breeds. In addition, humans and dogs share many similarities in disease development and expression, making dogs a valuable research model for human genetic traits and disease states. Most previous coat color genetic research has emphasized single gene traits, or simple interactions. However, as our understanding of genetics, pigmentation or otherwise, expands, the potential for complex gene interactions has been brought into the spotlight. An area of particularly promising potential is the *Agouti Signaling Protein* (*ASIP*) gene and identification of genes that can modify the phenotypes caused by various *ASIP* alleles. My research has revolved around the *ASIP* gene: characterizing the inheritance patterns of the allele series, identifying the causative mutation for the black-and-tan allele, raising questions regarding the postulated saddle tan allele, identifying a *Melanocortin 1 Receptor* (*MC1R*) allele that interacts with *ASIP*, and exploring the potential for modifier genes to alter the expected *ASIP* black-and-tan phenotype. With the knowledge that the wolf is the predecessor of the domestic dog (reviewed by Wayne and von Holdt 2012), including wolf samples in my research has allowed, in many instances, for the elucidation of wild-type alleles, speculation of the phenotypes present in early dogs, and the ability to better postulate the evolutionary origin of mutations.

2.0 LITERATURE REVIEW

2.1 Hair Growth and Pigment Production

Hair growth occurs in a follicle, which arises from a thickening of the ectoderm during embryologic development (Searle 1968). The follicle is a tubular shaped cavity located on top of a mesoderm papilla. The hair shaft is produced by a cell matrix, consisting of cells of the follicle, papilla and differentiated pigment cells (Searle 1968).

There are three main stages of the hair growth cycle, termed telogen, anagen, which is divided into 6 stages, and catogen, each of which has a distinct profile of melanocytes (Tobin 2008) (Figure 2.1) and melanogenesis-driving enzymes (Slominksi et al. 1994). During the rest stage or telogen, which occurs in most dog breeds but not in humans, no melanin production is apparent, although the components required for formation of a melanin-producing hair follicle are present (Tobin 2008). Commencement of the anagen stage marks the beginning of the growth cycle of the typical dog or mouse hair. Concentrations and activity of tyrosinase and dopachrome tautomerase, two components required for the production of melanin, increase dramatically in mouse hair follicles during this stage (Slominski et al. 1994). From anagen III to anagen VI there is an increase in the number of immature melanocytes differentiating to mature melanocytes, characterized by their increased size, development of a more extensive Golgi apparatus and endoplasmic reticulum, and increased dendricity (Tobin 2008). These mature melanocytes develop and accumulate in the hair bulb matrix, located above the follicular papilla, at the apex of the hair bulb (Tobin 2008). Mature melanocytes produce melanin pigments, which are accumulated in melanosomes, clustered into globules, excreted from the dendrites of the melanocyte, and taken up by phagocytosis by the keratinocytes of developing hair shafts (Ando et al. 2012).

Immature melanocytes exist in other areas surrounding the hair follicle, such as the outer root sheath and peripheral regions of the hair bulb, but these melanocytes are less differentiated and generally do not produce melanin (Tobin 2008). It is suggested that these immature melanocytes serve as a reservoir (Nishimura et al. 2002) for repopulating the collection of mature melanocytes that undergo apoptosis during the catagen, or regression, stage of hair growth (Tobin et al. 1998). During catagen, melanogenesis ceases prior to the cessation of hair

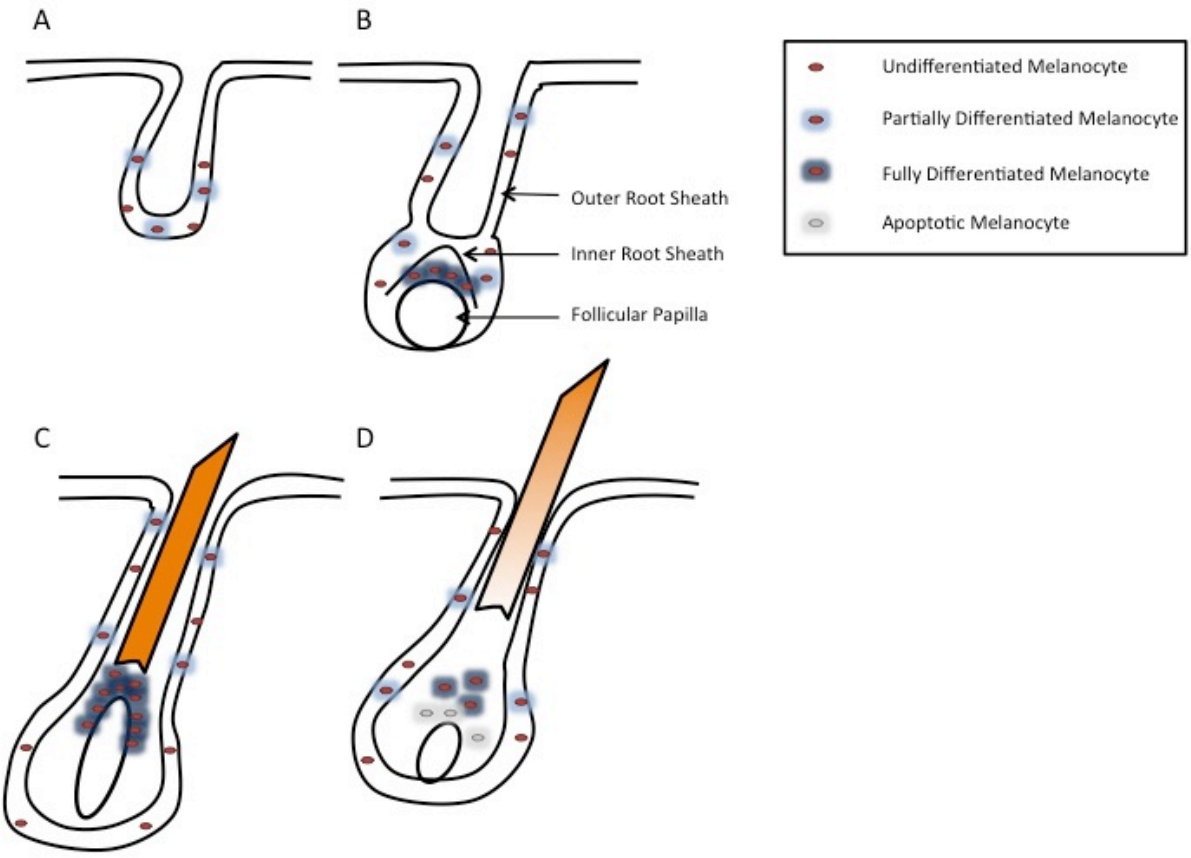


Figure 2.1: Melanogenesis during hair growth. A) Telogen. In the rest stage, undifferentiated and partially differentiated melanocytes are present in the outer root sheath. B) Anagen I-III. Melanocytes become differentiated and accumulate between the developing inner root sheath and follicular papilla. C) Anagen IV-VI. The primary stage of hair growth, with mature melanocytes supplying pigmented melanosomes to the developing hair shaft. D) Catagen. Melanogenesis ceases, melanocytes go through apoptosis, and hair growth terminates. (Adapted and redrawn from Tobin 2008)

growth (Tobin 2008) and is accompanied by dramatic decreases in tyrosinase and dopachrome tautomerase activity (Slominski et al. 1994).

Pigment production in mammalian cells takes place in melanosomes (Seiji et al. 1963), organelles specific to melanocytes. During fetal development, melanoblasts originate from the neural crest on the dorsal surface of the embryo and migrate down the body ventrally to their final destination (Jackson et al. 1994) where they mature into melanocytes. Two forms of

melanin can be produced by melanocytes: eumelanin and phaeomelanin. As outlined by Prota (1980), eumelanin is a black-based pigment produced by the conversion of tyrosine to 3,4-dihydroxyphenylalanine (dopa) and then dopaquinone through exposure to the enzyme, tyrosinase. Dopaquinone is cyclized to leucodopachrome, which is decarboxylized to 5,6-dihydroxyindole, and in turn oxidized to eumelanin. Conversely, in the presence of cysteine or glutathione, dopaquinone is converted to cysteinyldopa and ultimately to phaeomelanin, a red-based pigment (Prota 1980). Melanosomes produce either eumelanin or phaeomelanin, though melanocytes can have both eumelanin- and phaeomelanin-producing melanosomes (Tobin 2008). The variable expression of these two pigments, controlled by many different genes, produce the myriad of coat colors and patterns observed in mammalian species.

2.2 Coat Color Genetics

Since mammals are capable of producing only two pigment types, it was believed that the genetics of coat color production involved understanding the factors affecting the relative shape, size, number and position of granules containing these pigments, or the substitution of one pigment type for the other (Searle 1968). The ability for melanocytes to produce both eumelanin- and phaeomelanin-producing melanosomes allows for the expression of these two pigment types simultaneously or sequentially. Mutations in genes and promoters or regulatory elements also alter the temporal and spatial expression of these two pigments (Lu et al. 1994). The patterns in which the pigment types are expressed in dogs depend on a number of genes that encode for receptor proteins (Newton et al. 2000), ligands (Berryere et al. 2005), and transcription factors (Karlsson et al. 2007).

2.2.1 Melanocortin 1 Receptor

The *Melanocortin 1 Receptor* (*MC1R*) gene encodes for a 317 amino acid seven transmembrane domain G-protein coupled receptor located on the surface of melanocytes (Chhajlani et al. 1992; Robbins et al. 1993; Newton et al. 2000). Binding of the extracellular ligand α-Melanocyte Stimulating Hormone (α-MSH), to a functional MC1R protein results in an increased concentration of intracellular cAMP (Mountjoy et al. 1992; Ollmann et al. 1998), triggering the biochemical process that ultimately results in the production of eumelanin. If α-MSH is prevented from binding to MC1R, either through a mutation of *MC1R* or by antagonism by ASIP, the melanocyte instead produces phaeomelanin.

Probably the most widely recognized researcher of early dog coat color inheritance, C. C. Little (1957), predicted four alleles, reiterated by Willis (1989), at what Little termed the *E* locus. This is now mapped to *MC1R* (Newton et al. 2000) on dog chromosome 5. There are several alleles of *MC1R*, and resulting phenotypes, including: E^M mask, *E* wild-type, e^{br} brindle, and *e* red. Winge (1950) used different letters for alleles that caused similar sets of phenotypes, and in this case assigned A instead of E.

At the start of this study, three alleles had been confirmed for the *E* locus, or *MC1R* gene. The wild-type allele, *E*, allows for the production of both eumelanin and phaeomelanin in a pattern determined by the alleles of *Agouti Signaling Protein* (*ASIP*; Schmutz et al. 2003) and *Canine Beta Defensin 103* (*DEFB103*; Candille et al. 2007). A loss of function mutation, R306*, produces a premature stop codon and a truncated protein, rendering the MC1R receptor non-functional and preventing the production of eumelanin (Newton et al. 2000; Everts et al. 2000). Dogs homozygous for this allele, termed *e*, can only produce phaeomelanin pigment, so will therefore be solid red, yellow, or cream in color. Loss of function *MC1R* mutations that result in predominantly red phenotypes have been reported in numerous other species, including mice (Robbins et al. 1993), cattle (Klungland et al. 1995; Joerg et al. 1996), fox (Vage et al. 1997), pigs (Kijas et al. 2001), and horses (Marklund et al. 1996). The E^M allele, dominant to both *E* and *e*, produces a eumelanistic mask on the muzzle of the dog (Schmutz et al. 2003). This mask sometimes extends over much of the face, and may also result in eumelanin pigment on the top surface of the base of the tail. The E^M phenotype is most readily visible in dogs with an a^y fawn *ASIP* genotype, though an a^t black-and-tan genotype will allow for some eumelanin pigment to be visible in the cheek points that are generally pigmented with phaeomelanin. A fourth e^{br} allele, causing brindle, had been predicted by Little (1957) and Willis (1989), but was subsequently mapped to the *DEFB103* gene, termed the *K* locus (Kerns et al. 2007).

The majority of alleles associated with *MC1R* are related to the general ability of the melanocytes to produce either phaeomelanin or eumelanin. Predicted alleles of *MC1R* relate more to the overall expression of one or the other pigment, and less to patterns of these pigments. However, identification of the E^M allele, producing a eumelanistic mask, suggests that mutations within *MC1R* may also influence dispersion of eumelanin and phaeomelanin in predictable patterns.

2.2.2 Agouti Signaling Protein

Agouti Signaling Protein encodes a 131 amino acid paracrine-signaling molecule (Bultman et al. 1992). The ASIP protein is an extracellular ligand for MC1R and functions as an antagonist to α-Melanocyte Stimulating Hormone (α-MSH) (Bultman et al. 1992). Binding of ASIP to MC1R prevents the production of eumelanin pigment and allows for the production of phaeomelanin in a temporal or spatial manner (Lu et al. 1994). The alleles of *ASIP* dictate the pattern in which phaeomelanin is expressed.

Dog *ASIP* consists of three coding exons and at least one non-coding first exon (Kerns et al. 2004). Mice have alternate first exons for *ASIP* that produce various isoforms (Vrieling et al. 1994). Mouse exon 1A and 1A' are responsible for ventral phaeomelanin expression, while exon 1B and 1C are present in isoforms associated with production of alternating eumelanin and phaeomelanin bands along individual hair shafts (Vrieling et al. 1994). Since mice and dogs share many common coat color phenotypes, it is possible that dog *ASIP* has additional first exons, although they have not been identified at the present time.

Little (1957), postulated an allele series for what he termed the *A* locus, now mapped to *ASIP* on canine chromosome 24 (Kerns et al. 2004). Little's *A* locus consisted of four alleles: A^S solid color, a^y sable, a^w wild-type, and a^t black-and-tan. Seven years earlier, Winge (1950) postulated six alleles for an equivalent locus, which he named in terms of *C*: C^t solid color, *C* wild-type, c^{bi} black-and-tan, c^{br} brindle, c^{sa} saddle tan, and c^{ma} melanistic mask. Finally, Willis (1989) predicted a different six alleles for the *A* locus: *A* solid color, a^y sable, a^g wild-type, a^s saddle tan, a^t black-and-tan, and *a* recessive black. At the onset of this study, data for four *A* locus alleles: a^y sable, a^w wild-type, a^t black-and-tan, and *a* recessive black were available. The presence of an additional a^s saddle tan allele, was inconclusive. The causative mutations for the a^y fawn (Berryere et al. 2005) and *a* recessive black (Kerns et al. 2004) alleles had been identified in the coding region of *ASIP*.

The phenotypes caused by the postulated *ASIP* alleles are quite varied. The a^y allele produces a phenotype generally termed fawn, though known by other names such as sable or mahogany in some breeds. The predominant pigment in the fawn phenotype is phaeomelanin, though eumelanin can be present on the tips of longer hairs, or as a result of the E^M mask allele of *MC1R,* on the face and tail. The a^w wild-type allele, also termed sable or wolf-sable, is similar to what is observed in wild canids such as the wolf and coyote. The individual hairs are generally

banded, with the base and tip pigmented by eumelanin and a subapical band of phaeomelanin. Some solid colored black hairs are sometimes also present. The black-and-tan phenotype, associated with the a^t allele, is a predominantly eumelanistic phenotype, with phaeomelanin restricted to distinct regions on the lower limbs, cheeks, eyebrows, chest, and around the anus. The recessive black *a* allele produces a solid eumelanin phenotype (Kerns et al. 2004) and is primarily found in herding breeds (Berryere et al. 2005). There are a few isolated exceptions, such as Schipperkes and Samoyeds, where herding breeds may have contributed to the breed development at some point. The postulated a^s allele, associated with the saddle tan phenotype, produces a pattern whereby the limbs, head, and tail of the dog are pigmented with phaeomelanin and eumelanin pigment is restricted to a saddle-shaped patch on the back of the dog.

To date, two of the predicted *ASIP* alleles have been successfully identified in the coding region of *ASIP*. The recessive black allele, *a*, postulated by Willis (1989), is caused by an R96C amino acid substitution in exon 4 of *ASIP* (Kerns et al. 2004). The allele for fawn/sable, termed a^y by Little (1957) and Willis (1989), is caused by two adjacent amino acid substitutions, A82S R83H, also in exon 4 of *ASIP* (Berryere et al. 2005). The causative mutation for the solid color phenotype, assumed to be A^s by Little (1957), C^t by Winge (1950) and *A* by Willis (1989), has since been associated with a ΔG23 mutation in *Canine Beta Defensin 103* (*DEFB103*), now referred to as the K^B allele of the *K* locus (Candille et al. 2007). Likewise, the brindle phenotype, termed c^{br} and predicted to be part of the *ASIP* series by Winge (1950), has also been mapped to the *K* locus and renamed k^{br}, with the recessive allele of the K series, k^y, allowing for the expression of the *ASIP* phenotypes (Kerns et al. 2007). The melanistic mask phenotype, assumed to be the c^{ma} allele by Winge (1950), has since been associated with a M264V amino acid substitution in the *MC1R* gene and termed E^M (Schmutz et al. 2003).

At the onset of this research, some points that remained to be addressed regarding *ASIP* in dogs, were the identification of the genetic cause of the a^t black-and-tan allele, determining whether or not the a^s saddle tan allele is an allele of *ASIP* or the result of a modification by an interacting gene, and determining other potential instances of gene interaction resulting in *ASIP* allele modification.

2.2.3 Pigment Switching

As demonstrated by some of the *ASIP* alleles in mice (Vrieling et al. 1994), and long associated with dogs with various suspected *ASIP* mutants (Little 1957), eumelanin and

phaeomelanin can be alternately expressed along the hair shaft, resulting in a pattern of bands along the length of the hair. The exact mechanism behind this pigment-type switching has yet to be fully explained. Since phaeomelanin pigment production results from the binding of ASIP to MC1R and eumelanin pigment production results from the binding of α-MSH to MC1R, switching between the two pigment types indicates a transient, yet renewable, expression or activity of the two ligands. Production of phaeomelanin pigment has been associated with the level of *ASIP* mRNA present in mouse skin tissue (Bultman et al. 1992). Four alternate first exons of mouse *ASIP* have been identified, and their mRNA expression levels correlate with ventral production of phaeomelanin (exons 1A and 1A') or phaeomelanin production during the mid-phase of the anagen hair growth cycle (exons 1B and 1C) resulting in the characteristic hair-banding pattern (Vrieling et al 1994).

The switch from eumelanin to phaeomelanin production in dogs in accordance with the a^y, a^w, or a^t alleles of *ASIP* is prevented by the dominant K^B mutation in *DEFB103* (Candille et al. 2007). The K^B allele binds to MC1R and results in a solid eumelanin coat color by functioning as an antagonist for ASIP yet still allowing α-MSH to bind and trigger the production of eumelanin (Candille et al. 2007). It is apparent that the production of pigmentation patterns incorporating both eumelanin and phaeomelanin, at least in species such as dogs, rely on the interplay of *ASIP*, *MC1R*, and *DEFB103*.

2.2.4 Coat Color Phenotype Variation

There has been much advancement in our knowledge of dog coat color genetics since the early days of research that focused on observation of color outcomes from designed breeding trials. Dog coat colors and patterns can be simplified by considering a few classifications, such as which pigment types are present: only eumelanin or phaeomelanin or both; what pattern the pigments occur in; whether those pigments are diluted from their original form; and if there are areas that are white with no pigment at all. This series of questions can paint a reasonably accurate picture of any individual dog, both in phenotype and genotype. However, upon closer inspection, there are even more subtle differences within these broad classifications of color and pattern.

Focusing specifically on the alleles of *ASIP*, there appears to be variation among the accepted phenotypes for a^y fawn, a^w wolf sable, and a^t black-and-tan. There are at least three distinct phenotypes, termed grizzle/domino, saddle tan, and black-and-tan, that could all be

classified in the broader category of black-and-tan, having phaeomelanin on the limbs, ventral surface, and face, with eumelanin predominating on the torso (Figure 2.2). Likewise, within both the a^y fawn and a^w wolf sable phenotypes, there appears to be a vast difference in the amount of eumelanin expressed in relation to phaeomelanin, making certain dogs with either an $a^y/\_$ or $a^w/\_$ genotype appear darker than others with the same genotype. Breeder lore and anecdotal accounts have attempted to address these issues with varying degrees of success. From a genetics standpoint, the potential explanations include the presence of separate alleles that independently produce coincidentally similar phenotypes, or the function of additional genes that modify the action of existing coat color alleles, producing similar yet distinct phenotypes.

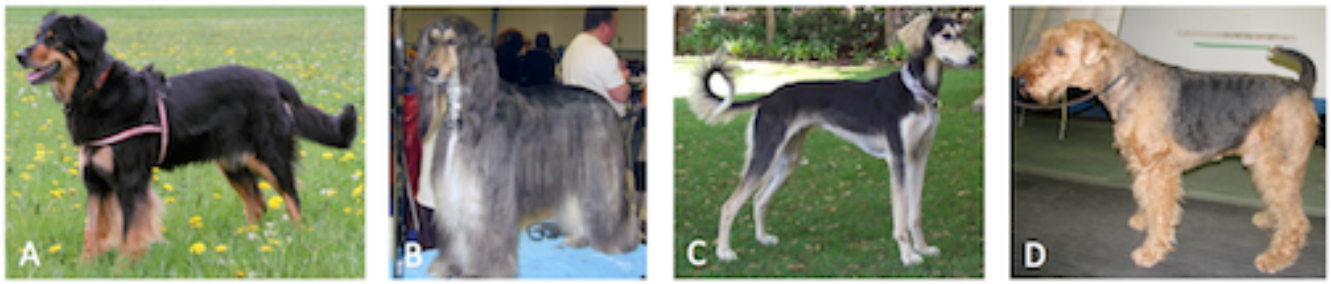


Figure 2.2: Four dog color patterns with phenotypic similarity in phaeomelanin pigment on the ventral surface and eumelanin pigment on the dorsal surface. A) A Hovawart with the black-and-tan phenotype. B) An Afghan Hound with the domino pattern. C) A Saluki with the grizzle pattern. D) An Airedale Terrier with the saddle tan pattern.

2.3 Short Interspersed Nuclear Elements (SINEs)

2.3.1 Transposable Elements

A number of types of transposable elements exist throughout mammalian genomes. They can be divided into two groups: DNA transposable elements, which make use of transposases to cut out a region of genomic DNA and reinsert it in a new location (Sinzelle et al. 2009), and retrotransposons, which involve the insertion of reverse transcribed RNA into genomic DNA (Belancio et al. 2009). Retrotransposons can be further divided into subcategories differentiated by the inclusion of long terminal repeats. The short retrotransposon subcategory, which does not incorporate long terminal repeats, include long interspersed elements (LINEs) and short

interspersed elements (SINEs) (Belancio et al. 2009). While LINEs have two open reading frames that encode proteins required for their transposition (Moran et al. 1996; Feng et al. 1996), SINEs do not include an open reading frame, therefore do not encode for any protein (reviewed in Belancio et al. 2008).

2.3.2 Canine SINEs

The major subfamily of SINE elements in dogs has been termed SINEC_Cf and has been reported to have a frequency of approximately 233,000 repeats in the CanSS Poodle assembly (Wang and Kirkness 2005). In addition, over 11,000 loci are reported to be bimorphic for SINEC_Cf elements, meaning that sequence fragments were identified to contain a SINEC_Cf element where no element exists in the CanSS Poodle assembly or CanFam1 Boxer assembly (Wang and Kirkness 2005). This indicates a potential for many unidentified SINEs to exist in the canine genome that cannot be predicted through investigation of publically available assemblies. While many SINEC_Cf elements are located in noncoding regions, approximately 40% of annotated canine genes containing at least one element (Wang and Kirkness 2005).

2.3.3 Impact of SINEs

SINEs can impact gene expression and phenotype development in a number of ways. While it can be assumed that the majority of SINEs are inserted in non-coding regions of the genome, this is certainly not the case in all instances. Exonization, interruption of coding regions, and recombination are the most common causes of SINEs introducing genomic instability. Human *Alu* elements, a primate-specific family of SINEs, are rarely reported in coding regions, though the interruption of coding sequence by SINEs is possible. An *Alu* element inserted in exon 2 of *human cholinesterase* (*ChE*) has been associated with a hereditary form of acholinesterasemia (Muratani et al. 1991). The inheritance of the *ChE Alu* element appears to function in a dose-dependent manner, with individuals with two copies of the insertion allele showing no *ChE* activity, and individuals heterozygous for the insertion allele having only half the regular level of *ChE* activity (Muratani et al. 1991). Some cases of Huntington Disease in humans have been associated with an *Alu* insertion within an intron of the *α-Adducin* gene (Goldberg et al. 1993). Both of these retrotransposition events have resulted in an *Alu* element in the sense direction, therefore not necessarily altering splice patterns due to incorporation of a new 3' acceptor site.

Alu elements can cause homologous recombination through crossover since they often appear in close proximity to one another, they have high sequence homology, and due to the high number of elements in the genome, increase the probability for recombination (Callinan and Batzer 2006). Callinan and Batzer (2006) report that from 1999 to 2006, 25 new *Alu-Alu* recombination events have been reported to cause human disease traits.

When present in the reverse, or anti-sense, orientation, human SINEs can potentially introduce a new splice site and can cause exonization of intronic sequence (Sorek et al. 2004). This same phenomenon is recognized in dogs where the canine SINEC_Cf contains sequence motifs that introduce a new 3' splice acceptor site (Wang and Kirkness 2005). As such, SINEs can alter splicing, incorporating the SINE sequence and novel sequence into a new coding exon. A database search by Krull et al. (2005) identified 153 loci that contain exons produced by *Alu* element exonization. These are cases of exonization that have become fixed in the primate population, and are not associated with disease states or phenotype alterations due to variability. As such, exonization by retrotransposable elements not only plays a role in production of some disease states, but has also contributed to the evolution and development of the primate genome.

2.3.4 Retrotransposons in Coat Color Genetics

In addition to affecting disease states, SINEs have been identified as causing pigment phenotypes in some species. A 198 bp SINE element has been identified approximately 3000 bp 5' of the canine *Micro-opthalmia-associated Transcription Factor* (*MITF*) melanocyte-specific first exon and is associated with white spotting patterns in some breeds (Karlsson et al. 2007). Further research with the *MITF* SINE characterized two separate modes of inheritance producing either a codominant or recessive piebald white phenotype, depending on dog breed (Schmutz et al. 2009). Since this SINE is present in the forward orientation, it is predicted that it may affect the transcription of the melanocyte-specific first exon, without effect on hearing or vision which are controlled in part by alternate first exons of *MITF* located further upstream (Schmutz et al. 2009).

The black-eyed white phenotype in mice has been attributed to a LINE insertion in intron 3 of *MITF* (Yajima et al. 1999). The presence of the transposable element disrupts the expression and splicing of the *MITF-A* and *MITF-H* isoforms, which results in lowered mRNA levels and an associated deafness phenotype (Yajima et al. 1999). The insertion also completely eliminates the

expression of the *MITF-M*, or melanocyte specific isoform, producing a solid white unpigmented phenotype (Yajima et al. 1999).

The merle phenotype in dogs, characterized by random patches of deeply pigmented and dilute hairs, is caused by a SINE insertion in intron 10 of *PMEL* (Clark et al. 2006). The merle SINE is present in the reverse orientation which introduces an alternate splice site in intron 10 (Clark et al. 2006). Merle shows a codominant inheritance pattern, with dogs heterozygous for the SINE insertion expressing the traditional merle phenotype. Homozygous dogs are predominantly white and have an increased risk of vision and hearing abnormalities.

In Normande cattle, the brindle phenotype has been attributed to an 8402 bp LINE insertion resulting in the overexpression of exon 2C of *ASIP* (Girardot et al. 2006). In the presence of an *E/E MC1R* genotype, the brindle A^{br} allele produces a dominant phenotype of alternating stripes of black and red pigment (Girardot et al. 2006).

2.4 Methods Used to Identify Coat Color Genes

2.4.1 Candidate Gene Analysis

When the genetic cause of a trait is known in one species and the associated gene and similar trait exist in another species, there is strong evidence to support investigation of the common gene as a cause of the common phenotype. The mouse model has proven to be very valuable in this pursuit. In terms of mammalian coat color genetics, the understanding of the function of the *MC1R* gene in mice suggested it as a candidate gene for the solid phaeomelanin phenotype in cattle (Joerg et al. 1996), horses (Marklund et al. 1996), and dogs (Newton et al. 2000), among others. Further, the use of analogous correlations of phenotype and causative genes across species has been utilized in furthering the understanding of disease states in humans where animal models allow for genetic selection and manipulation. For example, congenital ichthyosis, a skin condition, exists in both dogs and humans. Grall et al. (2012) identified the causative mutation for congenital ichthyosis in Golden Retrievers as an indel in the *patatin-like phospholipase domain-containing protein 1* (*PNPLA1*) gene. Further investigation of the homologous human gene identified a number of mutations in individuals affected with congenital ichthyosis (Grall et al. 2012).

2.4.2 Linkage

Linkage mapping relies on the process of recombination, whereby genetic material is exchanged between parental chromosomes in the offspring. An estimated rate of recombination is based on genetic distance between the genetic marker and the locus in question, and is equal to a 1% chance of recombination for a distance of 1 cM, or 1 linkage unit (Lewin 2000). This way, by genotyping related individuals for a series of genetic markers and analyzing the amount of recombination based on phenotype, regions of low recombination can be identified. These regions of low recombination suggest linkage with the causative mutation for the phenotype in question.

An example of a causative mutation being discovered through the use of linkage mapping is the K^B allele of the *K* locus in dogs (Kerns et al. 2007). Kerns et al. (2007) utilized linkage mapping to identify three loci with logarithm of odds (LOD) scores over 2 for the dominant black phenotype. Increasing the number of markers surrounding the key loci and analyzing for the occurrence of recombination between marker and locus recombination refined the candidate region to a region of 12 Mb on chromosome 16 (Kerns et al. 2007). Further fine mapping and association studies allowed this larger mapped region to be decreased to an area including 16 candidate genes, and additional sequencing identified the causative mutation for the K^B dominant black allele as a three base pair deletion in *DEFB103* (Candille et al. 2007).

2.4.3 Genome Wide Association Study (GWAS)

2.4.3.1 History of Genome Wide Association Studies

Identification of the causative mutation for any particular trait often relies on pursuing functional candidate genes or positional candidate genes. If a similar trait has been characterized in another species or a gene is known to function in a particular manner, then that gene may be explored in a novel species in hopes of identifying a causative mutation for the similar trait. This is known as a functional candidate gene, if the gene can also be implicated in being involved with the production of the trait in a new species. A trait is mapped to a particular chromosomal location through association of marker genotypes with the phenotype in a group of related individuals. Traditional linkage analysis methods involve a relatively small number of markers spread over a chromosome of interest or the entire genome. Through identifying recombination events between related individuals, regions of the genome can be eliminated from potentially

harboring a causative gene or mutation. The downfall of this method is that a small number of markers results in low mapping resolution, or difficulty in identifying a precise location for a potential causative mutation. More often than not, linkage mapping will identify a relatively broad region on a specific chromosome as being potentially important for the trait of interest, but a researcher will then have to shift to candidate gene identification to pinpoint the actual causative gene in that region or identify further polymorphic markers to conduct fine mapping of the candidate region. Identifying candidate genes relies on an assumption that the causative mutation is located within a gene and that the gene has been previously identified, or at least predicted. It is possible that linkage analysis or GWAS will identify a region of the genome where no genes are currently annotated, at which point identification of candidate genes becomes decidedly more difficult.

In 1996, Risch and Merikangas formulated a theory that suggested that association studies utilizing a large number of SNPs spread over the entire genome would be more statistically powerful and require a smaller sample size than the linkage analysis techniques. They predicted that this method would be particularly beneficial in cases of complex traits involving multiple gene interactions or contributions. The only hurdle to these association studies was predicted to be the technological ability to genotype a large number of SNPs, theoretically numbering around 1,000,000, in a feasible manner. Within ten years of the Risch and Merikangas (1996) publication, the sequencing of the human genome and development of SNP arrays were allowing for genome wide association studies (GWAS) for complex human disease traits (Visscher et al. 2012). While traditional linkage techniques are based on highly polymorphic markers, such as microsatellites, that are relatively widely spaced, GWAS utilizes bi-alleleic SNPs that are located much closer together. Since the first large scale published GWAS research in 2007 (Wellcome Trust Case Control Consortium 2007) GWASs have rapidly gained in popularity, from approximately 20 GWAS projects published within the first six months of 2007, to nearly 2,500 publications in the first six months of 2011 (Visscher et al. 2012).

2.4.4 GWAS Methods and Interpretation

GWAS utilizes the concept of linkage disequilibrium (LD) between genotyped common SNPs and ungenotyped causative mutations, given that the causative mutation is relatively common within the tested population (Visscher et al. 2012). As reviewed by McCarthy et al.

(2008), many of the early human GWAS experiments were designed with matching case and control samples. It was pointed out, however, that cohort population studies held significant promise for elucidating genetic contributions to continuous or quantitative traits (McCarthy et al. 2008). Individuals are genotyped for a large number of SNPs selected for coverage and linkage disequilibrium structure. These SNPs are arranged on a chip or array platform to facilitate ease and speed of genotyping. Arguably, the most array options are available for humans, with chips designed for variable levels of coverage and density. However, as genome sequencing projects advanced, SNP chips were designed for an increasing number of species (Table 2.1).

Table 2.1: SNP chips available through GeneSeek.

Species	Chip Name	No. of SNPs
Dog	CanineHD	170,000
Cattle	GoldenGate BeadChip	3072
Cattle	BovineHD	777,000
Cattle	Bovine SNP50 BeadChip	54,609
Pig	Porcine SNP60 BeadChip	62,163
Sheep	Ovine SNP50 BeadChip	54,241
Horse	Equine SNP50 BeadChip	54,602

There are three methods of visualizing data resulting from a GWAS analysis, as described by McCarthy et al. (2008) (Figure 2.3). (1) A quantile-quantile diagram plots a test statistic, commonly a chi-squared, for each SNP against the expected statistic value under the null hypothesis. Qualifying the deviation of observed values from the expected values can indicate an excess of disease associations within the population. (2) Signal intensity plots attempt to determine the diagnostic ability of individual SNPs. Genotype data is plotted against axes representing intensity values for each allele. The ability of data points to cluster according to their genotypes indicates the strength of the diagnostic abilities of the particular SNP for the trait of interest. (3) Likely the most common visualization tool for GWAS data is the Manhattan plot. The x axis represents the SNP location sequentially along the genome, while the y axis

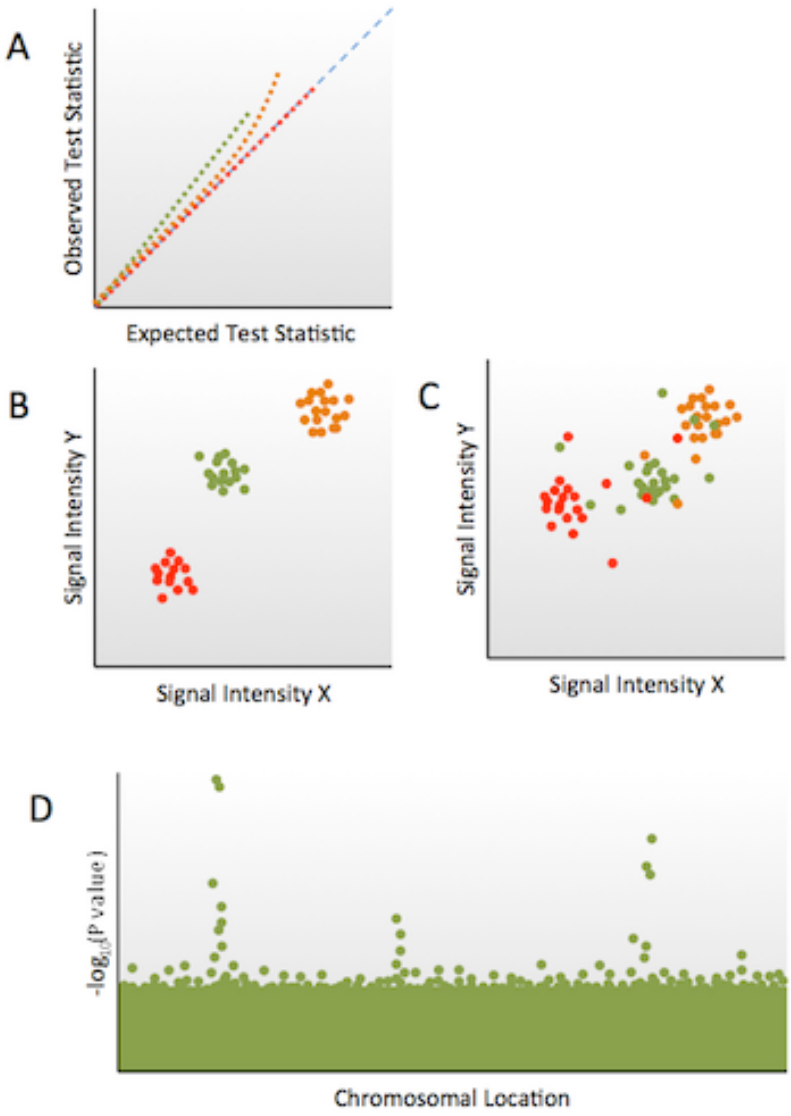


Figure 2.3: Three common visualization tools for GWAS data. A) A quantile-quantile plot graphs the difference of observed statistics from the expected (blue dashed line). The red dashed line demonstrates no deviation from the expected, therefore no association with the trait. The green dashed line shows proportional deviation from the expected line, suggesting cryptic relatedness. The orange dashed line shows strong evidence of association with the desired trait. B) and C) are cluster diagrams that attempt to cluster data points based on intensity. B) Data points group into observable genotype groupings. C) Data points do not group reliably with similar genotypes. D) A Manhattan graph plots data points for each GWAS SNP against the negative log of the P value for that point. Higher points indicate association of that chromosomal location with the trait of interest. (Adapted from McCarthy et al. 2008.)

represents the $-\log_{10}$(P value), quantifying the significance of a given SNP to the trait of interest. Depending on the goals of a given GWAS project, any or all of these interpretive methods may be used.

For complex traits, both quantitative and qualitative, the number of significant loci, generally deemed to have P values $< 5 \times 10^{-8}$, that have been identified increase with the sample size (Visscher et al. 2012). This outcome can be beneficial in cases where a large number of contributing genes are suspected, but confounding when attempting to identify a single primary causative gene.

2.4.5 Identification of Causative Mutations

Despite the potential for GWAS studies to elucidate influential genes associated with given complex traits, there have been relatively few causative variants for such traits identified through GWAS studies (Visscher et al. 2012). This may, in part, be due to the use of GWAS for investigation of increasingly complex disease traits resulting from multifactorial inheritance and environmental contributions. However, there have been a number of causative mutations identified through the application of GWAS methodologies to less complex traits, attributed to either single gene inheritance or interaction of a smaller number of contributing genes. Likewise, successful identification of traits specific to individual breeds or populations within a species has occurred with use of GWASs.

A series of GWAS experiments utilizing the Affymetrix v2.0 canine SNP chip, with 100K SNPs, on a selection of dogs from the CanMap data set identified mutations in each of three genes, *R-spondin-2* (*RSPO2*), *fibroblast growth factor-5* (*FGF5*) and *keratin-71* (*KRT71*) that are associated with coat length and type in dogs (Cadieu et al. 2009). The causative mutation for German White Fleckvieh syndrome, a condition affecting eye function and development, hypopigmentation and hearing loss in the German White Fleckvieh cattle breed, has been identified in the *MITF* gene after conducting a GWAS with the Illumina bovine high density 777K SNP chip (Philipp et al. 2011). Identification of a indel mutation in *patatin-like phospholipase domain-containing protein 1* (*PNPLA1*) through use of a GWAS in Golden Retriever dogs with congenital ichthyosis led to the subsequent identification of a causative mutation for a similar condition in humans (Grall et al. 2012). While many current GWAS studies culminate at the point of locus identification, these few examples demonstrate the

potential for GWASs to facilitate the discovery of causative mutations and assist in the advancement of research across species.

3.0 A SINE INSERTION CAUSES THE BLACK-AND-TAN AND SADDLE TAN PHENOTYPES IN DOMESTIC DOGS[1]

3.1 Abstract

Agouti Signaling Protein controls the localized expression of red and black pigment in the domestic dog through interaction with other genes, such as *Melanocortin 1 Receptor* and *Beta-Defensin 103*. Specific *ASIP* alleles are necessary for many of the coat color patterns, such as black-and-tan and saddle tan. Mutations in two *ASIP* alleles, a^y and *a*, have previously been identified. Here we characterize a mutation consisting of a SINE insertion in intron 1 of *ASIP* that allows for the differentiation of the a^w wolf sable and a^t black-and-tan alleles. The SINE insertion is present in dogs with the a^t and *a* alleles, but absent from dogs with the a^w and a^y alleles. Dogs with the saddle tan phenotype were all a^t/a^t and Schnauzers with the salt-and-pepper phenotype were all a^w/a^w. Genotypes of 201 dogs of 35 breeds suggest that there are only four *ASIP* alleles, as opposed to the five or six predicted in previous literature. These data demonstrate that the dominance hierarchy of *ASIP* is: $a^y > a^w > a^t > a$.

3.2 Introduction

Agouti Signaling Protein has been implicated in coat color production in a variety of domestic animals including: horses (Rieder et al. 2001), cattle (Girardot et al. 2005), pigs (Drögemüller et al. 2006), and dogs (Kerns et al. 2004, Berryere et al. 2005) (Appendix 3.1). *ASIP* alleles can act as an antagonist to α- MSH, preventing the binding of α-MSH to MC1R and resulting in the production of phaeomelanin over some or all of the body (Lu et al. 1994). The function of *ASIP*, and in turn the expression of phaeomelanin and eumelanin, is dependent on mutations in *ASIP*, as demonstrated by the nonagouti alleles causing solid black in sheep (Royo et al. 2008), dog (Kerns et al. 2004), and horse (Rieder et al. 2001). As demonstrated in mice (Vrieling et al. 1994), there can be multiple transcripts of *ASIP* resulting from alternate first exons.

Over the past sixty years, many authors have attempted to predict the various alleles of *ASIP*, or the *A* locus, in dogs. In 1950, Winge predicted six alleles at what he termed the *C* locus,

[1] A version of this chapter has been published: Dreger DL, Schmutz SM. 2011. A SINE insertion causes the black-and-tan and saddle tan phenotypes in domestic dogs. Journal of Heredity. 102:S11-S18.

and called them: solid color (C^t), wolf sable (C), black-and-tan (c^{bi}), brindle (c^{br}), saddle tan (c^{sa}), and melanistic mask (c^{ma}). Little, (1957) postulated four *A* locus alleles producing: solid color (A^s), fawn (a^y), wolf sable (a^w), and black-and-tan (a^t). Finally, Willis (1989) predicted six A locus alleles causing: solid color (A), fawn (a^y), wolf sable (a^g), saddle tan (a^s), black-and-tan (a^t), and recessive solid black (a). Modern nomenclature for dog coat color alleles is taken largely from the work of Little (1957). Recent advances in coat color genetic research have shown that some of the earlier predicted alleles are alleles at other genes. Brindle (k^{br}) and solid color (K^B) have been mapped to *DEFB103* (Kerns et al. 2007; Candille et al. 2007). Melanistic mask (E^M) is caused by a M264V mutation of *Melanocortin 1 Receptor* (Schmutz et al. 2003). The allele causing fawn (a^y), predicted as part of the *A* locus series by Little (1957) and Willis (1989) is caused by two point mutations in adjacent amino acids, A82S R83H, in exon 4 of *ASIP* (Berryere et al. 2005). The recessive solid black allele is caused by a R96C mutation in exon 4 of *ASIP* (Kerns et al. 2004), although this allele is rare. The predicted *ASIP* alleles that cannot currently be distinguished are wolf sable (a^w), black-and-tan (a^t), and saddle tan.

The a^w wolf sable allele causes a pattern of hairs that are banded in alternating sections of eumelanin and phaeomelanin along the hair shaft. Dogs that are a^w wolf sable have a predominantly phaeomelanin ventral surface, with banded hairs appearing on the dorsal surface and head (Figure 3.1a). The Keeshond pictured in Figure 3.1a has diluted phaeomelanin that appears cream, whereas in other breeds this pigment can be a deeper red/orange color. The black-and-tan phenotype (Figure 3.1b) is a very distinct pattern of phaeomelanin points on a eumelanin background. The "tan points" that are phaeomelanin in color are traditionally the distal portions of the legs, along the sides of the muzzle, small dots above each eye, on the caudal surface of the chest, and around the anus and ventral surface of the tail. The saddle tan phenotype (Figure 3.1e) is very similar to the black-and-tan phenotype, except the phaeomelanin points are larger in size, expanding completely up the legs onto the shoulder and hip, and over the entire face and head. Eumelanin pigmentation is restricted to a "saddle" shaped area on the dorsal surface.

Figure 3.1: Coat color variation attributed to the *ASIP* locus. A) a wolf sable (a^w) Keeshond, B) a black-and-tan (a^t) Doberman Pinscher, C) a fawn (a^y) Shetland Sheepdog, D) a recessive black (a) German Shepherd Dog, E) a saddle tan Basset Hound

3.3 Materials and Methods

3.3.1 Dogs and Families

Two hundred one dogs from 35 breeds were selected to investigate the interaction of the *ASIP* alleles and their associated phenotypes. The coat color patterns expressed include wolf sable (n = 57), fawn (n = 48), black-and-tan (n = 30), saddle tan (n = 8), solid eumelanin (n = 34), solid phaeomelanin (n = 11) and miscellaneous patterns (n = 13). Five breeds: Siberian Husky, Alaskan Malamute, Keeshond, Swedish Vallhund, and Norwegian Elkhound, were selected for study because they were suspected of being fixed at a^w/a^w, the wolf sable pattern. Ten breeds, Cardigan Welsh Corgi, Pembroke Welsh Corgi, Collie, Shetland Sheepdog, Eurasier, Dachshund, Saluki, German Shepherd Dog, Finnish Lapphund and Jack Russell Terrier, were chosen because they exhibit phenotypes attributed to multiple *ASIP* alleles. Five breeds that exhibit the saddle tan phenotype: Airedale Terrier, German Shepherd Dog, Beagle, Basset Hound, and Pembroke Welsh Corgi were included in the study. Thirty-six dogs from 12 hunting breeds, including Brittany Spaniel, French Brittany Spaniel, Chesapeake Bay Retriever, German Longhair Pointer, German Shorthair Pointer, German Wirehair Pointer, English Springer Spaniel, Golden Retriever, Large Munsterlander, Labrador Retriever, Weimaraner, and Vizsla, were included because these breeds commonly have solid coat colors caused by the K^B allele of *DEFB103* (Candille et al. 2007) or the *e/e* genotype of *MC1R* (Newton et al. 2000). These genotypes are epistatic to the *ASIP* alleles, so the *ASIP* genotypes of many hunting breeds were largely unknown. Thirteen additional dogs of six breeds were selected because they exhibit unusual color patterns for their breed. These include 2 fawn Akitas with more than the usual amount of eumelanin pigment, a Border Collie with banded hairs, a wild-boar Dachshund, a "patterned sable" Shar Pei, a salt-and-pepper Standard Schnauzer, and 7 Miniature Schnauzers, 3 of which are salt-and-pepper, 2 black-and-silver, and 2 solid black. DNA samples and plucked hair samples were collected from 10 Keeshonds with the wolf-sable pattern and 9 Shetland Sheepdogs with the fawn pattern. Hair samples were examined under a dissecting microscope to observe and record banding pattern and relative length of the hair with eumelanin pigment (Appendix 3.2 and Appendix 3.3). DNA samples were gathered using cheek brushes (Epicentre, Madison, WI) and were collected widely from Canada and the United States. The dogs utilized in this research were owned by private individuals and, in accordance with the

Canadian Animal Care guidelines, signed consent for use of the DNA in coat color studies was obtained from the owners.

3.3.2 Primer Design and PCR

An 35 kb region between the predicted exon 1 (Kerns et al. 2004) and the exon 2 start codon of *ASIP* was compared between two publicly available BAC sequences from a black-and-tan (a^t) Doberman Pinscher (GenBank AC092250) and a fawn (a^y) Boxer (GenBank NW_876277). Primers were subsequently designed for this region (Appendix 3.4).

Segments were amplified by polymerase chain reaction (PCR) in 15 µl reactions consisting of 1.5 µl 10x PCR Buffer (Fermentas), 0.3 µl of 10 mM dNTP, 0.9 µl of 25 mM $MgCl_2$, 10 pmol/µl each of forward and reverse primers, 0.1 µl of 5u/µl *Taq* polymerase (Fermentas), 9.2 µl dH_2O, and 1 µl of roughly 50 ng/µl DNA template. PCR was carried out in Stratagene Robocycler Gradient40 machines, with 4 minutes initial denaturation at 94°C, followed by 35-37 cycles of 50 second 94°C denaturation, 50 second annealing at primer specific temperatures, and 50 second 72°C extension. This was followed by a final 4 minute 72°C extension period. Product bands were excised from 2% agarose gel and isolated using the QIAquick gel extraction kit (Qiagen, Mississauga, ON), and sequenced at the National Research Council of Canada Plant Biotechnology Institute, using an ABI Prism 373 Sequencer (Perkin Elmer Corporation) and the Big Dye Terminator kit (Perkin Elmer Corporation). Sequences were aligned using the Sequencher 4.8 software program (Gene Codes Corporation, Ann Arbor, MI).

3.3.3 Genotyping

Dogs were genotyped for the a^y allele of *ASIP* as previously described (Berryere et al. 2005). A portion of the dogs were also genotyped for the *a* allele of *ASIP* (Kerns et al. 2004). Since the *a* allele is rare outside of herding breeds, only individuals suspected of carrying *a* were genotyped. Dogs were also genotyped for the presence of a newly discovered SINE insertion, described below.

3.4 Results

3.4.1 Sequence Analysis

A polymorphism consisting of a 239 bp SINE insertion was found in the Doberman Pinscher, but not in the Boxer. A second SINE is located in both Boxer and Doberman Pinscher

23.5 kb upstream of the exon 2 start codon and approximately 215 bp downstream of the variable SINE.

The fixed SINE, found in the forward orientation, is present in dogs of all *ASIP* genotypes. The more 5' variable SINE is in the reverse orientation, and is present only in dogs with a^t or *a ASIP* alleles. The 215 bp region between these two SINEs is found in dogs of all *ASIP* genotypes, as are flanking regions 3' of the fixed SINE and 5' of the variable SINE (Figure 3.2) (GenBank HQ910236, HQ910237, HQ910238, HQ910239) (Appendix 3.5).

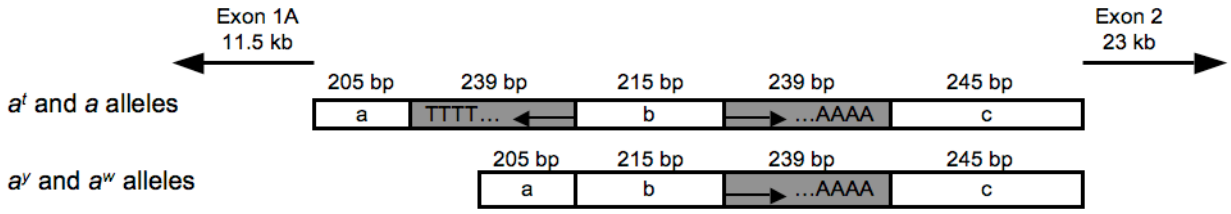


Figure 3.2: An illustration of an 1143 bp fragment, from 5' of the *ASIP* start codon in exon 2, obtained from genomic DNA from a black-and-tan (a^t) Shetland Sheepdog and a wolf sable (a^w) Siberian Husky. Shaded boxes represent SINEs, with arrows denoting orientation. Segments a, b, and c of this fragment are present in all dogs.

The two SINE insertions are 95% similar to each other on average, excluding variability in the length of the poly-A tail. The fixed SINE is 94.8% similar to the family of canine SINEs (SINEC_Cf) described by Wang and Kirkness (2005), while the variable SINE is 92.3% similar to the Wang and Kirkness SINEC_Cf sequence. Variants within the SINEs, and between the fixed and variable SINEs, included single nucleotide polymorphisms and single nucleotide insertions/deletions. None of these polymorphisms appeared to affect the pigmentation phenotypes.

3.4.2 Individual Genotypes

Two hundred one dogs were genotyped for the *ASIP* alleles (Table 3.1 and Table 3.2), including the variable SINE insertion. All a^t black-and-tan (n = 30) and *a* recessive black (n = 9) dogs had only the larger fragment which included the variable SINE insertion. The dogs that are a^y fawn (n = 48) and a^w wolf sable (n = 57) all had at least one copy of the smaller fragment

which lacks the variable SINE insertion. The subset of 46 heterozygous dogs with both the larger and smaller fragments were either fawn or wolf sable in color (Table 3.1), with the exception of four hunting breed dogs which were solid black due to the K^B allele of *DEFB103* or solid red due to an *e/e* genotype at *MC1R*. All 8 dogs described as saddle tan were determined to be a^t/a^t by genotyping for the SINE insertion (Table 3.1). This genotype therefore occurs in dogs that are traditional black-and-tan and also in dogs that are saddle tan. All 8 Schnauzers, whether their phenotype was salt-and-pepper, black-and-silver or black were genotyped for the SINE insertion and found to be a^w/a^w wolf sable (Table 3.1).

Dogs of many sporting breeds exhibit coat colors of either solid phaeomelanin or solid eumelanin, with or without white spotting. These phenotypes are caused by an *e/e* genotype at *MC1R* (Newton et al. 2000) or the K^B allele of *DEFB103* (Candille et al. 2007). These genotypes are epistatic to the *ASIP* alleles. Thirty-six dogs belonging to 12 sporting breeds were genotyped for the variable SINE insertion and the fawn a^y allele. Since the *a* allele is not expected to occur in sporting breeds, it was not genotyped in these dogs. Three alleles, a^t, a^y, and a^w, were present in these breeds (Table 3.2).

Of the 35 breeds genotyped, the 5 that were predicted to be fixed for a^w: Siberian Husky, Alaskan Malamute, Swedish Vallhund, Norwegian Elkhound, and Keeshond, were indeed found to be homozygous for absence of the variable SINE (Table 3.1). After excluding dogs that genotyped as having the a^y allele, twelve additional breeds were found to have one allele that did not have the variable SINE. Only one breed, the Eurasier, was found to have all four *ASIP* alleles.

3.4.3 Dominance hierarchy of ASIP alleles

The dominance hierarchy of the four *ASIP* alleles can be determined by observing the genotype results, relative to the phenotypes observed. A number of heterozygous dogs were genotyped, including 4 a^y/a^w, 28 a^y/a^t, 4 a^y/a, 11 a^w/a^t, 3 a^w/a and 3 a^t/a dogs (Table 3.1 and Table 3.2). Forty-seven of these dogs do not have the $K^B/\_$ or *e/e* genotypes that are epistatic to the *ASIP* alleles, so were used to determine *ASIP* allele inheritance. All dogs with the a^y allele

Table 3.1: *ASIP* genotypes of 166 dogs of 23 breeds illustrating that the variable SINE is present in both the a^t and a allele.

Breed	Coat Color	n	*ASIP* Genotype	Presence of Variable SINE	A82S[a]	R96C[b]
a^w/a^w expected						
Keeshond	Wolf sable	11	a^w/a^w	No	A/A	R/R
Alaskan Malamute	Wolf sable	12	a^w/a^w	No	A/A	R/R
Norwegian Elkhound	Wolf sable	11	a^w/a^w	No	A/A	R/R
Siberian Husky	Wolf sable	7	a^w/a^w	No	A/A	R/R
Swedish Vallhund	Wolf sable	5	a^w/a^w	No	A/A	R/R
Multiple ASIP alleles expected						
Basset Hound	Solid red	1	a^t/a^t	Yes	A/A	R/R
	Fawn	3	a^y/a^t	Yes	S/A	R/R
Cardigan Welsh Corgi	Black and tan	2	a^t/a^t	Yes	A/A	R/R
	Fawn	2	a^y/a^y	No	S/S	R/R
	Fawn	1	a^y/a^t	Yes	S/A	R/R
Collie	Black and tan	3	a^t/a^t	Yes	A/A	R/R
	Fawn	2	a^y/a^y	No	S/S	R/R
	Fawn	5	a^y/a^t	Yes	S/A	R/R
Dachshund	Black and tan	4	a^t/a^t	Yes	A/A	R/R
	Fawn	1	a^y/a^y	No	S/S	R/R
	Fawn	2	a^y/a^t	Yes	S/A	R/R
Eurasier	Black	5	a/a	Yes	A/A	C/C
	Wolf sable	1	a^w/a^w	No	A/A	R/R
	Wolf sable	1	a^w/a^t	Yes	A/A	R/R
	Wolf sable	1	a^w/a	Yes	A/A	C/R
	Fawn	1	a^y/a^w	No	S/A	R/R
	Fawn	3	a^y/a^t	Yes	S/A	R/R
	Fawn	1	a^y/a	Yes	S/A	C/R

Table 3.1: Continued

Breed	Coat Color	n	*ASIP* Genotype	Presence of Variable SINE	A82S[a]	R96C[b]
Finnish Lapphund	Black and tan	2	a^t/a^t	Yes	A/A	R/R
	Fawn	3	a^y/a^t	Yes	S/A	R/R
German Shepherd Dog	Black	3	a/a	Yes	A/A	C/C
	Black and tan	7	a^t/a^t	Yes	A/A	R/R
	Black and tan	2	a^t/a	Yes	A/A	C/R
	Wolf sable	6	a^w/a^t	Yes	A/A	R/R
	Wolf sable	2	a^w/a	Yes	A/A	C/R
Jack Russell Terrier	Black and tan	2	a^t/a^t	Yes	A/A	R/R
	Fawn	2	a^y/a^t	Yes	S/A	R/R
Pembroke Welsh Corgi	Fawn	1	a^y/a^t	Yes	S/A	R/R
Saluki	Black and tan	2	a^t/a^t	Yes	A/A	R/R
	Brown and tan	1	a^t/a^t	Yes	A/A	R/R
	Fawn	2	a^y/a^y	No	S/S	R/R
	Fawn	6	a^y/a^t	Yes	S/A	R/R
Shetland Sheepdog	Black	1	a/a	Yes	A/A	C/C
	Black and tan	3	a^t/a^t	Yes	A/A	R/R
	Black and tan	1	a^t/a	Yes	A/A	C/R
	Fawn	2	a^y/a	Yes	S/A	C/R
	Fawn	2	a^y/a^t	Yes	S/A	R/R
	Fawn	9	a^y/a^y	No	S/S	R/R
Dogs of miscellaneous colors						
Akita	Dark Fawn	1	a^y/a^w	No	S/A	R/R
	Dark Fawn	1	a^y/a^y	No	S/S	R/R
Border Collie	Banded hairs	1	a^w/a^w	No	A/A	R/R
Dachshund	Wild-boar	1	a^w/a^t	Yes	A/A	R/R
Miniature Schnauzer	Black	2	a^w/a^w	No	A/A	R/R
	Black and silver	2	a^w/a^w	No	A/A	R/R
	Salt and pepper	3	a^w/a^w	No	A/A	R/R

Table 3.1: Continued

Breed	**Coat Color**	**n**	***ASIP* Genotype**	**Presence of Variable SINE**	**A82S**[a]	**R96C**[b]
Standard Schnauzer	Salt and pepper	1	a^w/a^w	No	A/A	R/R
Shar Pei	Patterned sable	1	a^w/a^w	No	A/A	R/R
Saddle tan pattern						
Airedale Terrier	Saddle tan	2	a^t/a^t	Yes	A/A	R/R
Basset Hound	Saddle tan	1	a^t/a^t	Yes	A/A	R/R
Beagle	Saddle tan	2	a^t/a^t	Yes	A/A	R/R
German Shepherd Dog	Saddle tan	1	a^t/a^t	Yes	A/A	R/R
Pembroke Welsh Corgi	Saddle tan	2	a^t/a^t	Yes	A/A	R/R

[a] 82S is indicative of the a^y allele

[b] 96C is indicative of the *a* allele

expressed a fawn phenotype, regardless of the second allele present, indicating that this is the top dominant allele of the hierarchy. Dogs with an a^w/a^w, a^w/a^t, or a^w/a genotype, all exhibited a wolf sable coat pattern, suggesting that a^w is dominant to a^t and a. Dogs with either an a^t/a^t or a^t/a genotype were black-and-tan. Only dogs with an a/a genotype expressed the recessive solid black phenotype. Therefore, the *ASIP* allele hierarchy was determined to be: $a^y > a^w > a^t > a$.

Table 3.2: *ASIP* genotypes illustrating that the a^t, a^y, and a^w alleles are present in sporting breeds.

Breed	**Coat Color**	**n**	***ASIP* Genotype**
French Brittany Spaniel	Black and white	1	a^y/a^t
Brittany Spaniel	Red and White	1	a^w/a^t
Chesapeake Bay Retriever	Brown	3	a^y/a^y
English Springer Spaniel	Black and white	1	a^t/a^t
German Longhair Pointer	Brown and white	4	a^t/a^t
German Shorthair Pointer	Brown and white	3	a^t/a^t
	Brown and white	1	a^w/a^t
German Wirehair Pointer	Brown and white	2	a^t/a^t
	Brown and white	1	a^w/a^w
Golden Retriever	Yellow	1	a^y/a^w
	Yellow	3	a^t/a^t
	Yellow	1	a^y/a^w
Labrador Retriever	Black	1	a^t/a^t
	Chocolate	2	a^t/a^t
	Yellow	2	a^t/a^t
Large Munsterlander	Black and white	3	a^t/a^t
Vizsla	Red	1	a^t/a^t
	Red	1	a^y/a^w
Weimaraner	Grey	4	a^t/a^t

3.5 Discussion

A black-and-tan phenotype is seen in mice and is caused by a 6 kb insertion in intron 1a of *ASIP* (Bultman et al. 1994). Similarly, in the dogs in this study with a black-and-tan phenotype or saddle tan phenotype an insertion of approximately 239 bp was found in intron 1a. In dogs, the insertion is a SINE.

Previous research has shown that some dog coat color patterns are caused by SINE insertions. A family of similar SINES, termed SINEC_Cf, is widely found throughout the canine genome (Wang and Kirkness 2005). A SINE insertion at the intron 10/exon 11 boundary of *SILV* causes the merle phenotype in dogs (Clark et al. 2006). The *SILV* SINE insertion, present in the reverse orientation, is predicted to alter exon splicing or lariat branching within intron 10 (Clark et al. 2006), as presence of a SINEC_Cf in the reverse orientation can introduce a 3' splice acceptor site (Wang and Kirkness 2005). A SINE insertion in the forward orientation, located 3.5 kb upstream of the *MITF* 1M promoter, causes the solid white coloring in breeds such as Boxers and Bull Terriers (Karlsson et al. 2007). This same insertion causes piebald spotting and pseudo-Irish spotting in numerous other breeds (Schmutz et al. 2009). Schmutz et al. (2009) predicted that the *MITF* SINE, present in the forward orientation, may alter transcription or regulatory elements required for expression of *MITF*.

Multiple alternate first exons are responsible for variable transcripts of *ASIP* in mouse. Vrieling et al. (1994) report four alternate first exons in mice, two of which, 1A and 1A' are associated with ventral specific production of phaeomelanin. The remaining two, 1B and 1C, are associated with hair cycle specific production of phaeomelanin, resulting in banded hairs. To date only one first exon has been described in dogs (Kerns et al. 2004). This first exon is orthologous to exon 1A of mouse. It was previously described in fox (Vage et al. 1997). Individual hairs of dogs that have an a^w or a^y allele have alternating bands of phaeomelanin and eumelanin along the hair shaft (Appendices 3.3 and 3.4). Individual hairs of dogs that have only the a^t or *a* alleles do not alternate between pigment types. Both have only solid eumelanin hairs on the dorsal area of the torso. Both *ASIP* SINEs are located 3' of the dog exon 1A, between exon 1A and the exon 2 start codon (Figure 3.2). Since the forward orientation fixed SINE is present with all *ASIP* phenotypes, it probably does not affect pigmentation. The location of the reverse orientation variable SINE, present in only the a^t and *a* alleles, may function to disrupt splicing of *ASIP* transcripts due to the 3' splice acceptor site described by Wang and Kirkness

(2005). Considering the mouse first exons described by Vrieling et al. (1994), the orthologous dog exon 1A is located upstream of the SINE insertions. The dog exon 1A, expected to affect ventral expression of *ASIP* in mice (Vrieling et al. 1994), is not impacted by the SINE insertions as dogs with the a^w, a^y, and a^t alleles express tan points. The mouse exons 1B and 1C, located between 1A and exon 2 and not yet identified in the dog, are responsible for hair cycle specific expression of *ASIP* (Vrieling et al. 1994), causing banding on the individual hairs. If regulatory exons orthologous to either mouse 1B or 1C exist in dogs, then the variable SINE insertion may disrupt those exons, preventing hair banding in dogs with the a^t and *a* alleles. A reverse variable SINE inserted prior to a regulatory exon, such as 1B or 1C, may cause alternate splicing of the exon, incorporating some of the SINE sequence into the regulatory exon. A disruption such as this in the regulatory exon responsible for producing phaeomelanin banding on individual hairs may prevent the *ASIP* expression required for such banding. This would be consistent with the a^t and *a* phenotypes seen in dogs. This suspected mode of action suggests that a regulatory exon may be located in close proximity to the variable SINE insertion. This provides information that may be valuable in identifying alternate first exons for dog *ASIP*. Retrotransposons have been shown to affect coat color phenotypes in dogs (Clark et al. 2006; Karlsson et al. 2007; Schmutz et al. 2009), but it is also possible that the *ASIP* SINE is in linkage disequilibrium with another polymorphism that is causative for the black-and-tan phenotype.

Utilizing formulas and methods suggested previously (Li et al. 1981; Yasue & Wada 1996; Chou et al. 2002), the divergence between the *ASIP* fixed SINE and the variable SINE is too great to suggest that the variable SINE is a direct duplication and inversion of the fixed SINE. I suggest that the variable SINE, present with the a^t and *a* alleles, resulted from a separate insertion event of a related SINE not identical to the fixed SINE already present in this region all dogs.

Different breeds have different sets of *ASIP* alleles. Utilizing data collected in the course of this research, as well as previously published a^y data (Berryere et al. 2005), and breed study data (Schmutz, unpublished), the *ASIP* alleles present in the 35 breeds examined in this study could be ascertained (Table 3.3). Despite the presumption that it is the ancestral *ASIP* allele of dogs, only 7 breeds appear to be fixed for the a^w allele. This suggests that there has been a high level of divergence through mutation and selection at the *ASIP* locus in domestic dog breeds. The a^w allele was found in only an additional 11 breeds, as one of multiple *ASIP* alleles present. The

a^t allele was found in 25 of the breeds studied. The a^y allele was found in 17 breeds in this study, in addition to another 16 breeds discussed by Berryere et al. (2005), but not utilized here.

The Fédération Cynologique Internationale (FCI) (2010) is the primary world federation of purebred dog registries. It was formed in 1911 and governs purebred standards for 84 member countries, recognizing 339 dog breeds. The FCI recognizes ten breed types: herders, pinschers/molossoid breeds, terriers, dachshunds, spitz, scent hounds, pointers, retrievers, companions, and sight hounds. Based on these breed type categories, the a^w allele was found in spitz, herder, pincher/molossoid, dachshund, retriever, and pointer groups (Table 3.3). The a^y allele was found in a wider selection of groups, including the pointer, retriever, sight hound, scent hound, pinscher/molossoid, herder, dachshund, terrier and spitz groups. The a^t allele was also found in a wide range of breed types, including the pointer, retriever, sight hound, scent hound, terrier, herder, dachshund, and spitz groups. The *a* allele was found only in herding breeds and one spitz breed in this study, though it has been documented in Samoyeds (Schmutz, unpublished data), another spitz breed, and more herding breeds not listed here (Berryere et al. 2005; Kerns et al. 2004). The narrow range of breed types in which the *a* allele is found in may suggest that it is a relatively recent mutation. The spitz breeds have been shown to be some of the earliest developed breeds (Parker et al. 2004). The high frequency of the a^w allele in many of the spitz breeds supports that the a^w allele is the ancestral *ASIP* allele in domestic dogs, despite the relatively small number of breeds with the allele.

While the a^w allele can be considered the ancestral *ASIP* allele, we suggest that the a^t allele is the result of a relatively old mutation as well. The a^t allele is present in a wide range of breed types, indicating its occurrence prior to modern breed development.

Only one breed, the Eurasier, was found to have all four *ASIP* alleles (Table 3.3). The Eurasier is a relatively recent breed developed in the 1960's from the Wolfspitz, commonly known as the Keeshond, the Chow Chow, and the Samoyed. This accounts for the presence of the a^w allele from the Keeshond, the a^y allele from the Chow Chow, and the *a* allele from the Samoyed. The presence of the a^t allele in Eurasiers is less certain, though it may have arisen from either the Wolfspitz or Chow Chow, prior to current breed standards that prohibit the black-and-tan pattern in those breeds. The breeding history of the Eurasier is reasonably well documented (Schneider and Schneider) and publicly accessible online pedigrees (http://www.berndschneider100845.de/) show the emergence of black-and-tan individuals from

Table 3.3: *ASIP* alleles present in 35 dog breeds. Breed type is classified based on the Fédération Cynologique Internationale standards.

Breed	**FCI Type**	***ASIP* Alleles**			
		a^y	a^w	a^t	a
Eurasier	Spitz	✓	✓	✓	✓
Border Collie	Herder		✓	✓	✓
German Shepherd Dog	Herder		✓	✓	✓
Dachshund	Dachshund	✓	✓	✓	
Golden Retriever	Retriever	✓	✓	✓	
Vizsla	Pointer	✓	✓	✓	
German Shorthair Pointer	Pointer		✓	✓	
German Wirehair Pointer	Pointer		✓	✓	
Brittany Spaniel	Pointer		✓	✓	
Akita	Spitz	✓	✓		
Shar Pei	Pinscher/Molossoid	✓	✓		
Keeshond	Spitz		✓		
Alaskan Malamute	Spitz		✓		
Norwegian Elkhound	Spitz		✓		
Siberian Husky	Spitz		✓		
Swedish Vallhund	Spitz		✓		
Standard Schnauzer	Pinscher/Molossoid		✓		
Miniature Schnauzer	Pinscher/Molossoid		✓		
Shetland Sheepdog	Herder	✓		✓	✓
Collie	Herder	✓		✓	
Cardigan Welsh Corgi	Herder	✓		✓	
Pembroke Welsh Corgi	Herder	✓		✓	
Finnish Lapphund	Spitz	✓		✓	
Jack Russell Terrier	Terrier	✓		✓	
Airedale Terrier	Terrier			✓	
Saluki	Sighthound	✓		✓	

Table 3.3: Continued

Breed	FCI Type	*ASIP* Alleles			
		a^y	a^w	a^t	a
Basset Hound	Scenthound	✓		✓	
French Brittany Spaniel	Pointer	✓		✓	
Large Munsterlander	Pointer	✓		✓	
Chesapeake Bay Retriever	Retriever	✓			
Labrador Retriever	Retriever			✓	
Beagle	Scenthound			✓	
English Springer Spaniel	Retriever			✓	
German Longhair Pointer	Pointer			✓	
Weimaraner	Pointer			✓	

reported matings between Chow Chows and Wolfspitz. Unfortunately, coat color is not documented for all individuals, so we are unable to determine whether the black-and-tan coloration was introduced by the Chow Chow or the Wolfspitz.

Based on the discovery of the variable SINE in this study, which can be used to identify the a^t allele, all four of the alleles at *ASIP* can now be discriminated with DNA testing. The present data (Table 3.1) demonstrate that dogs with the saddle tan phenotype are a^t/a^t at *ASIP*. The difference between the traditional black-and-tan pattern and the saddle tan pattern is likely caused by a modifier gene that allows for the expansion of the phaeomelanin points. A similar phenomenon involving expansion of phaeomelanin regions in black-and-tan mice has been observed and shown to be caused by a deletion of 216 kb, spanning from intron 1 of *TBX15* to 148 kb downstream of the adjacent *mannose-6-phosphate receptor* pseudogene (Candille et al. 2004). The saddle tan phenotype in mice, referred to as droopy ear, occurs in conjunction with craniofacial abnormalities and decreased body size (Candille et al. 2004). These conditions are not present with the saddle tan phenotype in dogs, suggesting that a more minor mutation within this region may only affect pigmentation, without the additional skeletal malformations. Somewhat surprisingly, the three phenotypes present in Miniature and Standard Schnauzers: salt-and-pepper, black-and-silver, and solid black; were all found to be a^w/a^w at *ASIP*. Little (1957) predicted that the salt-and-pepper phenotype of Schnauzers may be caused by the a^w allele,

though this suggestion has been discounted by many Schnauzer breeders. The solid black Schnauzers also have a K^B allele of *DEFB103* (unpublished data), which is epistatic to the *ASIP* alleles. Considering the relatively rare occurrence of the a^w allele in dogs, this finding may play a role in understanding Schnauzer breed development and history.

The limited occurrence of the a^w allele in breeds such as Shar-Pei, Border Collie, and Akita (Table 3.3), where it was unexpected, may be due to lack of selection against this color pattern, or the high frequency of the dominant a^y allele in the Shar-Pei and Akita. Most Shar-Peis and Akitas have only the dominant a^y allele at *ASIP* (unpublished data), so the rare occurrence of the recessive a^w allele may remain hidden for generations. The rare phenotype of a^w/a^w dogs in these breeds may be misclassified as fawn due to the similarity of these phenotypes in breeds with shorter coat length. The presence of the a^w allele in sporting breeds, such as pointers and retrievers, was unexpected. Though since the K^B allele, fixed in most of these breeds, is epistatic to the *ASIP* alleles, it is likely that the a^w allele has persisted in the sporting breeds due to lack of selection against it.

The presence of the variable SINE insertion in both the a^t and a alleles provides some insight into the evolution of the *ASIP* alleles (Figure 3.3). The a^t allele does not have any coding region polymorphisms when compared to the a^w allele. The only genetic difference between the a^w and a^t alleles is the presence of the variable SINE insertion in a non-coding region of a^t. This suggests that the a^t allele arose from the a^w allele and not from a^y. Additionally, the a allele does not have the a^y polymorphisms, but it does have the variable SINE insertion of a^t. This indicates that the a allele arose from the a^t allele, and differs from it solely by the point mutation in exon 4 (Kerns et al. 2004). The a^y allele arose independently from the a^w allele, through two point mutations in exon 4 (Berryere et al. 2005).

The characterization of the mutation responsible for the black-and-tan phenotype in dogs can lead to a better understanding of *ASIP* and how it interacts with *MC1R* and other modifier genes to produce the phenotypic variation in color seen in domestic dogs This research will continue with investigation of the SINE region of *ASIP* and its role in alternate transcript production. Exploration of potential modifier genes of *ASIP* may explain the variation seen within *ASIP* alleles, such as a^t/a^t saddle tan and the grizzle/domino phenotype in Afghan Hounds and Salukis that results from an interaction of *MC1R* and the a^t/a^t *ASIP* genotype (Dreger et al. 2010).

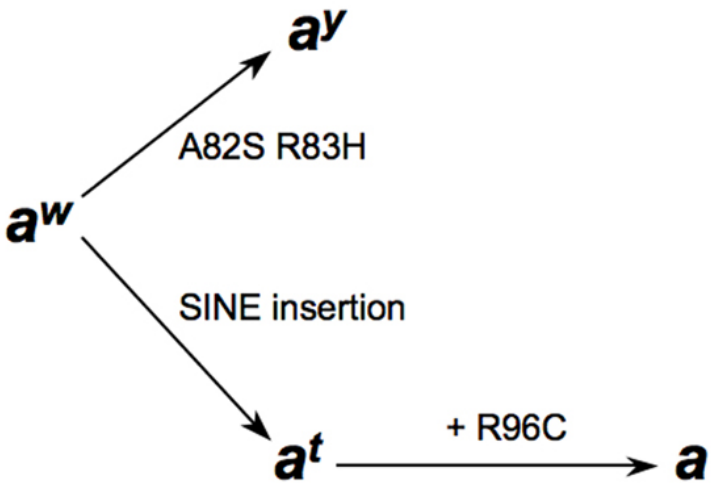


Figure 3.3: The evolution of the *ASIP* alleles. The a^y and a^t alleles arose independently from the a^w allele. The a allele arose from the a^t allele.

4.0 A NEW MUTATION IN MC1R EXPLAINS A COAT COLOR PHENOTYPE IN TWO "OLD" BREEDS: SALUKI AND AFGHAN HOUND[2]

4.1 Abstract

Melanocortin 1 Receptor has been studied in a wide variety of domestic animals (Klungland et al. 1995; Marklund et al. 1996; Vage et al. 1997; Kijas et al. 1998; Newton et al. 2000; Vage et al. 2003), and also several wild animals (Robbins et al. 1993; Ritland et al. 2001; Eizirik et al. 2003; Nachman et al. 2003; McRobie et al. 2009) in relation to coat color variation. A variety of phenotypic changes have been reported including coat colors from pure black to pure red, as well as some phenotypes with hairs with red and black bands. One phenotype, called grizzle in Salukis and domino in Afghan Hounds, appears to be unique to these 2 old dog breeds. This pattern is characterized by a pale face with a widow's peak above the eyes. The body hairs on the dorsal surface of Salukis and Afghan Hounds have both phaeomelanin and eumelanin portions, even through they had an a^t/a^t genotype at *ASIP*. In addition, all had at least one copy of a newly identified mutation in *MC1R*, g.233G>T, resulting in p.Gly78Val. This new allele, that we suggest be designated as E^G, is dominant to the *E* and *e* (p.Arg306ter) alleles at *MC1R* but recessive to the E^M (p.Met264Val) allele. The K^B allele (p.Gly23del) at *DEFB103* and the a^y allele (p.Ala82Ser and p.Arg83His) of *ASIP* are epistatic to grizzle and domino.

4.2 Introduction

The Saluki and Afghan Hound are both considered ''old'' breeds based on several genetic studies (Parker et al. 2004 ; Boyko et al. 2009) and are also closely related. Both Salukis and Afghan Hounds occur in a wide variety of coat colors (Table 4.1). Over the course of time, different colors have been selected more in one breed than the other and so the allele frequencies are not the same in both breeds. Both breeds share a coat color pattern that has not previously been studied using molecular genetics. This pattern is now known as domino in Afghans, after a famous show dog of that name. The pattern is called grizzle in Salukis, presumably because the dorsal surface looks like it contains a mix of eumelanin and phaeomelanin. In the course of a coat color study of Salukis, we became interested in the genes involved in these patterns, as well as their inheritance, which has been poorly understood.

[2] A version of this chapter has been published: Dreger DL, Schmutz SM. 2010. A new mutation in *MC1R* explains a coat color phenotype in 2 "old" breeds: Saluki and Afghan Hound. Journal of Heredity. 101:644-649.

Table 4.1: Previously described genotypes and phenotypes in Afghan Hounds and/or Salukis.

Color Phenotype	*MC1R*	*ASIP*	*DEFB103*
Fawn, masked	$E^M/\_$	$a^y/\_$	k^y/k^y
Fawn, no mask	*E*/*E* or *E*/*e*	$a^y/\_$	k^y/k^y
Black, brown or blue[a]	$E/\_$ or $E^M/\_$	any	$K^B/\_$
Red, yellow	*e*/*e*	any	any
Eumelanin-and-tan, masked	$E^M/\_$	a^t/a^t	k^y/k^y
Eumelanin-and-tan, no mask	$E/\_$	a^t/a^t	k^y/k^y
Brindle (full body)	$E/\_$ or $E^M/\_$	$a^y/\_$	k^{br}/k^{br} or k^{br}/k^y
Black-and-brindle	$E/\_$ or $E^M/\_$	a^t/a^t	k^{br}/k^{br} or k^{br}/k^y

[a] Most breeds are black because of K^B but herding breeds are often black because of *a*/*a*.

[b] The genotypes of cream dogs are not yet known, but an epistatic interaction occurs which seems to affect several genes (Schmutz, unpublished data).

4.3 Materials and Methods

4.3.1 Dogs and Families

DNA was obtained for this study from a number of Salukis and Afghan Hounds of various coat colors, including grizzle or domino pattern, using cheek brushes (Epicentre, Madison, WI). The Salukis were collected widely, from several countries. The Afghan Hounds were collected only in North America and a proportion of them represent a large multigeneration family that helped to verify the inheritance pattern. In addition, DNA from our bank of dogs studied previously in other coat color studies was utilized. Photographs of the dogs studied were used to verify the coat color and pattern. Hair samples were also examined in some cases. The dogs are owned by private individuals and signed consent from the owners was obtained for use of their DNA in coat color studies, in keeping with Canadian Animal Care guidelines.

4.3.2 Hair Analysis

Hair samples were pulled from the hip and shoulder area of some dogs that were grizzle or domino pattern. Individual hairs were examined under a dissecting microscope to confirm the root end and to check the distribution of phaeomelanin and eumelanin color.

4.3.3 DNA Sequencing

Because *MC1R* is a single exon gene, we were able to obtain the entire coding sequence from genomic DNA. Primers D and E (Newton et al. 2000) were used for amplification of a 1262-bp fragment. Polymerase chain reaction products were isolated using the QIAquick method (Qiagen, Mississauga, ON) and were sequenced at the National Research Council of Canada Plant Biotechnology Institute using an ABI Prism 373 Sequencer (Perkin Elmer Corporation) and the Big Dye Terminator kit (Perkin Elmer Corporation). Sequences were aligned using the Sequencher 4.8 computer software (Gene Codes Corporation, Ann Arbor, MI).

4.3.4 Genotyping

All individuals were genotyped for the presence of the a^y allele of *ASIP* using our previously published method (Berryere et al. 2005). A proportion of individuals were also genotyped for the *a*, or recessive black, allele of *ASIP* (Kerns et al. 2004), although this allele is atypical outside of herding breeds. Previously reported protocols were used to detect the *e* (Schmutz et al. 2002) and E^M alleles (Schmutz et al. 2003) of *MC1R*. A protocol was designed to detect the new *MC1R* mutation.

4.4 Results

4.4.1 Hair Analysis

The hairs from grizzle Salukis were typically a pale phaeomelanin shade at the base or root end and a dark eumelanin color at the tip. The proportions of phaeomelanin and eumelanin varied in different areas of the body but no hairs had alternating bands of phaeomelanin and eumelanin. There were some guard hairs that were entirely a eumelanin color. The shorter or undercoat hairs were a pale cream color. The hairs from domino Afghans were much longer and softer and had alternating bands of phaeomelanin and eumelanin of varying size on the same individual and among individuals. They also had some hairs with a phaeomelanin base and eumelanin at the terminal end.

4.4.2 MC1R Sequence Analysis

A previously unreported mutation was discovered in the *MC1R* sequence of several Salukis. The mutation is a g.233G**.** T substitution, resulting in an amino acid change from a glycine to valine at residue 78 (p.Gly78Val) (GenBank GU220379) which we have named E^G . Dogs that were heterozygous and homozygous for valine were identified. This mutation was observed only in Salukis (GenBank GU220378) and Afghan Hounds (GenBank GU220379) (Table 4.2) and not in any dogs of other breeds that were examined including Malamute, Shetland Sheepdog, and Siberian Husky (Appendix 3.5).

Table 4.2: *MC1R* and *ASIP* genotypes of Salukis and Afghan Hounds

Dog	**Breed**	**Phenotype**	***MC1R* Genotype**	***ASIP* Genotype**
Taz	Saluki	grizzle	E^G/E^G	a^t/a^t
Glitz	Saluki	grizzle	E^G/E^G	a^t/a^t
Flame	Saluki	grizzle	E^G/E^G	a^t/a^t
Azimuth	Saluki	grizzle	E^G/E^G	a^t/a^t
Seger	Saluki	grizzle	E^G/E^G	a^t/a^t
Lovey	Saluki	grizzle	E^G/E^G	a^t/a^t
Sue	Saluki	grizzle	E^G/E^G	a^t/a^t
Mooge	Saluki	grizzle	E^G/E^G	a^t/a^t
Fannie	Saluki	grizzle	E^G/E^G	a^t/a^t
Djinn	Saluki	grizzle	E^G/E^G	a^t/a^t
Mia	Saluki	grizzle	E^G/E^G	a^t/a^t
Portia	Saluki	grizzle	E^G/E^G	a^t/a^t
Cheyenne	Saluki	grizzle	E^G/E^G	a^t/a^t
Sarah Jane	Saluki	grizzle	E^G/E^G	a^t/a^t
Mariah	Saluki	grizzle	E^G/E	a^t/a^t
Cherokee	Saluki	grizzle	E^G/E	a^t/a^t
Tuvok	Saluki	grizzle	E^G/E	a^t/a^t
Zaynah	Saluki	grizzle	E^G/e	a^t/a^t

Table 4.2: Continued

Dog	Breed	Phenotype	*MC1R* Genotype	*ASIP* Genotype
Spirit	Saluki	grizzle	E^G/e	a^t/a^t
Faruza	Saluki	grizzle	E^G/e	a^t/a^t
Tyghar	Saluki	grizzle, brindle	E^G/e	a^t/a^t
Uno	Saluki	grizzle	E^G/e	a^t/a^t
Missy	Saluki	grizzle	E^G/e	a^t/a^t
Eli	Saluki	grizzle	E^G/e	a^t/a^t
Promise	Saluki	grizzle	E^G/e	a^t/a^t
Missie	Saluki	grizzle	E^G/e	a^t/a^t
Eden	Saluki	grizzle	E^G/e	a^t/a^t
Frank	Saluki	black-and-tan	E^M/E^G	a^t/a^t
Smurf	Saluki	black-and-tan	E/E	a^t/a^t
Falcor	Saluki	black-and-tan	E/E	a^t/a^t
Love	Saluki	black-and-tan	E/E	a^t/a^t
Scorch	Saluki	black-and-tan	E/E	a^t/a^t
Draco	Saluki	black-and-tan	E/E	a^t/a^t
Gladyr	Saluki	black-and-tan	E/E	a^t/a^t
Neptune	Saluki	black-and-tan	E/E	a^t/a^t
Saba	Saluki	black-and-tan	E/E	a^t/a^t
Onyx	Saluki	black-and-tan	E/E	a^t/a^t
Chester	Saluki	black-and-tan	E/E	a^t/a^t
Freckles	Saluki	black-and-tan	E/E	a^t/a^t
Flirt	Saluki	brown-and-tan	E/E	a^t/a^t
Chocolate	Saluki	brown-and-tan	E/E	a^t/a^t
Gala	Saluki	black-and-tan	E/e	a^t/a^t
Oasis	Saluki	black-and-tan	E/e	a^t/a^t
Dancer	Saluki	black-and-tan	E/e	a^t/a^t
Suzie	Saluki	black-and-tan	E/e	a^t/a^t
Ziva	Saluki	black-and-tan	E/e	a^t/a^t
Ricky	Saluki	black-and-tan	E/e	a^t/a^t

Table 4.2: Continued

Dog	Breed	Phenotype	*MC1R* Genotype	*ASIP* Genotype
Otis	Saluki	black-and-tan	E/e	a^t/a^t
Bertha	Saluki	cream	E^G/E^G	a^y/a^t
Sweetie	Saluki	cream	E^G/E^G	a^y/a^t
Farasha	Saluki	black	E^M/E^G	a^y/a^t
Chief	Saluki	fawn	E^G/e	a^y/a^t
Zia	Saluki	fawn	E^G/e	a^y/a^t
Simba	Saluki	fawn	E^G/e	a^y/a^t
Ali	Saluki	fawn	E/E	a^y/a^t
Lola	Saluki	fawn	E/E	a^y/a^t
Gabriel	Saluki	fawn	E/e	a^y/a^t
Jarel	Saluki	fawn	E/e	a^y/a^t
Ginger	Saluki	pale gold	E/e	a^y/a^y
Tess	Saluki	cream	E/e	a^y/a^t
Rasho	Saluki	black	E^M/e	a^y/a^t
Tirgan	Saluki	brindle, mask	E^M/e	a^y/a^y
Torc	Saluki	pale gold	e/e	a^y/a^y
Cygfa	Saluki	cream	e/e	a^y/a^y
Songbird	Saluki	cream	e/e	a^t/a^t
Millie	Saluki	cream	e/e	a^t/a^t
Zara	Saluki	cream	e/e	a^t/a^t
Boz	Afghan Hound	domino	E^G/E^G	a^t/a^t
Nola	Afghan Hound	domino, brindle	E^G/E^G	a^t/a^t
Weasel	Afghan Hound	domino	E^G/E^G	a^t/a^t
Blitz	Afghan Hound	domino	E^G/E^G	a^t/a^t
Speedy	Afghan Hound	domino	E^G/E^G	a^t/a^t
Issie	Afghan Hound	domino	E^G/E^G	a^t/a^t
Gio	Afghan Hound	domino	E^G/E^G	a^t/a^t
Inara	Afghan Hound	domino	E^G/E^G	a^t/a^t
Uma	Afghan Hound	domino, brindle	E^G/e	a^t/a^t

Table 4.2: Continued

Dog	Breed	Phenotype	*MC1R* Genotype	*ASIP* Genotype
Char	Afghan Hound	domino	E^G/e	a^t/a^t
Sara	Afghan Hound	domino	E^G/e	a^t/a^t
Beck	Afghan Hound	domino	E^G/e	a^t/a^t
Whineneit	Afghan Hound	domino	E^G/e	a^t/a^t
Shrek	Afghan Hound	domino	E^G/e	a^t/a^t
Goldie	Afghan Hound	domino	E^G/e	a^t/a^t
Icie	Afghan Hound	domino	E^G/e	a^t/a^t
Artie	Afghan Hound	domino	E^G/e	a^t/a^t
Kirt	Afghan Hound	domino	E^G/e	a^t/a^t
Jack	Afghan Hound	domino	E^G/e	a^t/a^t
Samantia	Afghan Hound	domino	E^G/e	a^t/a^t
Jayde	Afghan Hound	black-and-brindle	E^M/E^G	a^t/a^t
Zara	Afghan Hound	black-and-tan	E^M/E^G	a^t/a^t
Allie	Afghan Hound	black-and-tan	E^M/E^G	a^t/a^t
Scandal	Afghan Hound	black-and-brindle	E^M/E^G	a^t/a^t
Hector	Afghan Hound	black-and-tan	E^M/E^G	a^t/a^t
Prada	Afghan Hound	black-and-tan	E^M/E^G	a^t/a^t
Dude	Afghan Hound	black-and-tan	E^M/E^G	a^t/a^t
Drew	Afghan Hound	black-and-tan	E^M/E^G	a^t/a^t
Vanessa	Afghan Hound	black-and-brindle	E^M/E^G	a^t/a^t
Pristine	Afghan Hound	black-and-brindle	E^M/E^G	a^t/a^t
Memphis	Afghan Hound	black-and-tan	E^M/E	a^t/a^t
Skylar	Afghan Hound	black-and-brindle	E^M/e	a^t/a^t
Deedra	Afghan Hound	black-and-tan	E^M/e	a^t/a^t
Tarquin	Afghan Hound	black-and-tan	E^M/e	a^t/a^t
Andy	Afghan Hound	cream	E^G/E^G	a^y/a^t
Max	Afghan Hound	cream	E^G/E^G	a^y/a^t
Timer	Afghan Hound	cream	E^G/E^G	a^y/a^t
Fatima	Afghan Hound	fawn, mask	E^M/E^G	a^y/a^t

Table 4.2: Continued

Dog	Breed	Phenotype	*MC1R* Genotype	*ASIP* Genotype
Fire	Afghan Hound	fawn, mask	E^M/E^G	a^y/a^t
Axel	Afghan Hound	fawn, mask	E^M/E^G	a^y/a^t
Anastasia	Afghan Hound	fawn, mask	E^M/e	a^y/a^t
Azizah	Afghan Hound	fawn, mask	E^M/e	a^y/a^t
Raja	Afghan Hound	fawn, mask	E^M/E^M	a^y/a^t

Other mutations were also detected in the *MC1R* sequence. Some, such as p.Arg306ter, known as the *e* allele (Newton et al. 2000) and p.Met264Val, known as the E^M allele (Schmutz et al. 2003) were observed (Table 4.2). Other previously reported substitutions, not associated with a particular coat color phenotype, such as g.627C>T (Newton et al. 2000) were also observed. In addition some SNPs, not previously reported or available in GenBank sequence, were observed. These included g.–45A>G in the 5' region (GenBank GU233654) and g.*184G>C in the 3' region (GenBank GU233655), as well as a g.476C>A which resulted in a proline to glutamine change in residue 159 (p.Pro159Gln) (GenBank GU233656). All 3 of these variants were detected in Malamute and Siberian Husky, as well as in Saluki and Afghan Hound (Table 4.3). Haplotype analysis showed that the -45A allele was only present in dogs that also had a copy of 159Q. However, not all dogs homozygous for 159Q were homozygous for -45A. These 2 alleles may be in linkage disequilibrium (LD) but they are not in complete LD. None of the previously reported mutations or the newly discovered variants were in complete LD with the 78V allele (Table 4.3).

4.4.3 Phenotypic Association

The 2 coordinators of a recent study we conducted on coat color in Saluki, Margot Keast and Casey Gonda, developed the following definition: ''Grizzle—A pattern that is present at birth and remains throughout life, which comprises a dark overlay covering the top and sides of the body and outside of the limbs, from the top of the head to the tip of the tail. The darker color on the head gives the impression of a 'widow's peak' between the eyes where it contrasts with the lighter color of the face. The underside of the dog and the inner sides of the legs are the lighter base color. The overlay color may be black, gray/silver, red, or chocolate and can vary in

Table 4.3: Genotypes of SNPs, novel and previously reported, in Salukis and Afghan Hounds of various coat colors, and Shetland Sheepdogs, Malamutes, and Siberian Huskies with facial markings similar to grizzle/domino. SNP headings in italics were described prior to this research. Red letters represent the variant alleles at each SNP.

Name	Breed	Color	*MC1R* Genotype	*-45A>G*	G78V V = E^G	S90G	A104T	*P159Q*	*627C>T*	*M264V* V = E^M	H270D	Q278K	R302C	*R306ter* ter = T= *e*	**77C>G*	**184G>C*
Taz	Saluki	grizzle	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	G/G
Glitz	Saluki	grizzle	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	G/G
Flame	Saluki	grizzle	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	
Bertha	Saluki	cream	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	G/G
Sweetie	Saluki	cream piebald	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	G/G
Chief	Saluki	fawn	E^G/e	A/A	G/V	S/S	A/A	P/Q	C/C	M/M	H/H	Q/Q	R/R	C/T	C/C	G/G
Farasha	Saluki	black	E^M/E^G	G/G	G/V	S/S	A/A	P/P	C/C	M/V	H/H	Q/Q	R/R	C/C	C/C	G/C
Rasho	Saluki	black	E^M/e	G/A	G/G	S/S	A/A	P/Q	C/C	M/V	H/H	Q/Q	R/R	C/T	C/C	G/C
Tirgan	Saluki	brindle, mask	E^M/e	G/A	G/G	S/S	A/A	P/Q	C/C	M/V	H/H	Q/Q	R/R	C/T	C/C	G/C
Mr. Smurf	Saluki	black-and-tan	*E/E*	G/G	G/G	S/S	A/A	P/P	T/T	M/M	H/H	Q/Q	R/R	C/C	C/C	C/C
Ali	Saluki	fawn	*E/E*	G/G	G/G	S/S	A/A	P/P	T/T	M/M	H/H	Q/Q	R/R	C/C	C/C	C/C
Lola	Saluki	fawn	*E/E*	G/G	G/G	S/S	A/A	P/P	T/T	M/M	H/H	Q/Q	R/R	C/C	C/C	C/C
Songbird	Saluki	cream	*e/e*	A/A	G/G	S/S	A/A	Q/Q	C/C	M/M	H/H	Q/Q	R/R	T/T	C/C	G/G
Boz	Afghan Hound	domino	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	G/G
Nola	Afghan Hound	domino, brindle	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	G/G
Weasel	Afghan Hound	domino	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	G/G
Timer	Afghan Hound	cream	E^G/E^G	G/G	V/V	S/S	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	C/C	G/G
Uma	Afghan Hound	domino, brindle	E^G/e	G/A	G/V	S/S	A/A	P/Q	C/C	M/M	H/H	Q/Q	R/R	C/T	C/C	G/G
Char	Afghan Hound	domino	E^G/e	G/A	G/V	S/S	A/A	P/Q	C/C	M/M	H/H	Q/Q	R/R	C/T	C/C	G/G
Scandal	Afghan Hound	black-and-brindle	E^M/E^G	G/G	G/V	S/S	A/A	P/P	C/C	M/V	H/H	Q/Q	R/R	C/C	C/C	G/C
Allie	Afghan Hound	black-and-tan	E^M/E^G	G/G	G/V	S/S	A/A	P/P	C/C	M/V	H/H	Q/Q	R/R	C/C	C/C	G/C
Memphis	Afghan Hound	black and tan	E^M/E	G/G	G/G	S/S	A/A	P/Q	C/C	M/V	H/H	Q/Q	R/R	C/C	C/C	G/C
Nadia	Shetland Sheepdog	fawn/sable	*E/E*	G/G	G/G	G/G	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	G/G	G/G
Delainey	Shetland Sheepdog	fawn/sable	*E/E*	G/G	G/G	G/G	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	G/G	G/G
Akina	Shetland Sheepdog	fawn/sable	*E/E*	G/G	G/G	G/G	A/A	P/P	C/C	M/M	H/H	Q/Q	R/R	C/C	G/G	G/G
Dulli	Siberian Husky	wolf sable	*E/E*	G/A	G/G	G/G	T/T	P/Q	C/C	M/M	H/D	Q/K	C/C	C/C	C/C	C/C
Hank	Siberian Husky	wolf sable	*E/E*	G/G	G/G	G/G	T/T	P/P	C/C	M/M	D/D	K/K	C/C	C/C	C/C	C/C
Darby	Siberian Husky	wolf sable	*E/E*	A/A	G/G	G/G	T/T	P/Q	C/C	M/M	H/D	Q/K	C/C	C/C	C/C	C/C
Solara	Malamute	wolf sable	*E/E*	G/G	G/G	G/G	T/T	Q/Q	C/C	M/M	D/D	Q/Q	C/C	C/C	C/C	G/G
Blue	Malamute	wolf sable	*E/E*	G/G	G/G	G/G	T/T	P/Q	C/C	M/M	H/D	Q/K	C/C	C/C	C/C	G/C

intensity with the seasons and age. The base color may be various shades of red, fawn, golden, cream, or silver.'' The common characteristic that is most easily seen in grizzle Salukis (Figure 4.1) and domino Afghan Hounds (Figure 4.2) is the pale face with a pronounced widow's peak pattern above the eyes. We suggest that the new 78V allele be called E^G, consistent with other allele terminology used by dog breeders and owners since the classic work of Little (1957) who referred to this phenotype as grizzle in both Afghan Hounds and Salukis.

Figure 4.1: Photos of Salukis representing the grizzle phenotype (A and B) and the eumelanin-and-tan phenotype in both the brown-and-tan (C) and black-and-tan (D) variations.

Figure 4.2: Photos of Afghan Hounds with the domino phenotype (A and B) and the eumelanin-and-tan phenotype (C and D).

Not all the dogs with a 78V, or the E^G allele, had a grizzle or domino phenotype. Only dogs that also had an a^t/a^t genotype at *ASIP* had this phenotype. A single E^G allele in the presence of an *E* or *e* allele, in combination with an a^t/a^t genotype was sufficient to result in a grizzle or domino phenotype (Table 4.1), except in black dogs with a K^B allele. This suggests that the E^G allele is dominant to both the *E* and *e* allele. Both Saluki and Afghan Hounds

commonly have an a^y allele at *ASIP*, which is dominant to the a^t allele, and this a^y allele is epistatic to grizzle and domino. Neither the a^w nor the *a* allele are present in either breed (Dreger DL, Schmutz SM, Unpublished data).

Several Afghan Hounds we studied had an E^M allele and the E^G allele (Table 4.2). These dogs had a melanistic mask over most of their muzzle. If they were also a^t/a^t, they had tan or pale eyebrow spots. Their body color was also consistent with a black-and-tan pattern (Figure 4.2). Therefore, this suggests that the E^M allele is dominant to the E^G allele. A single Saluki had both an E^M allele and E^G allele and it also had a black-and-tan pattern on its body.

Dogs that have a K^B allele at *DEFB103* (Candille et al. 2007) are black, brown, or gray. This allele is epistatic to the alleles of *ASIP* and such black dogs did not show a grizzle pattern (Table 4.2). The K^B allele is relatively common in Afghan Hounds and relatively rare in Salukis. However, a few black dogs in each breed were included in this study.

4.4.4 Family Studies

A large multigeneration family of Afghan Hounds that included dogs that were domino, black-and-tan, fawn/sable, and cream was also studied. As predicted by the results in the individual Salukis, the inheritance of domino requires that both an a^t/a^t genotype at *ASIP* and the presence of at least one copy of the 78V mutation, or E^G allele, and no E^M or K^B allele are present (Figure 4.3). Therefore, the inheritance pattern does not follow a simple single gene mode of inheritance.

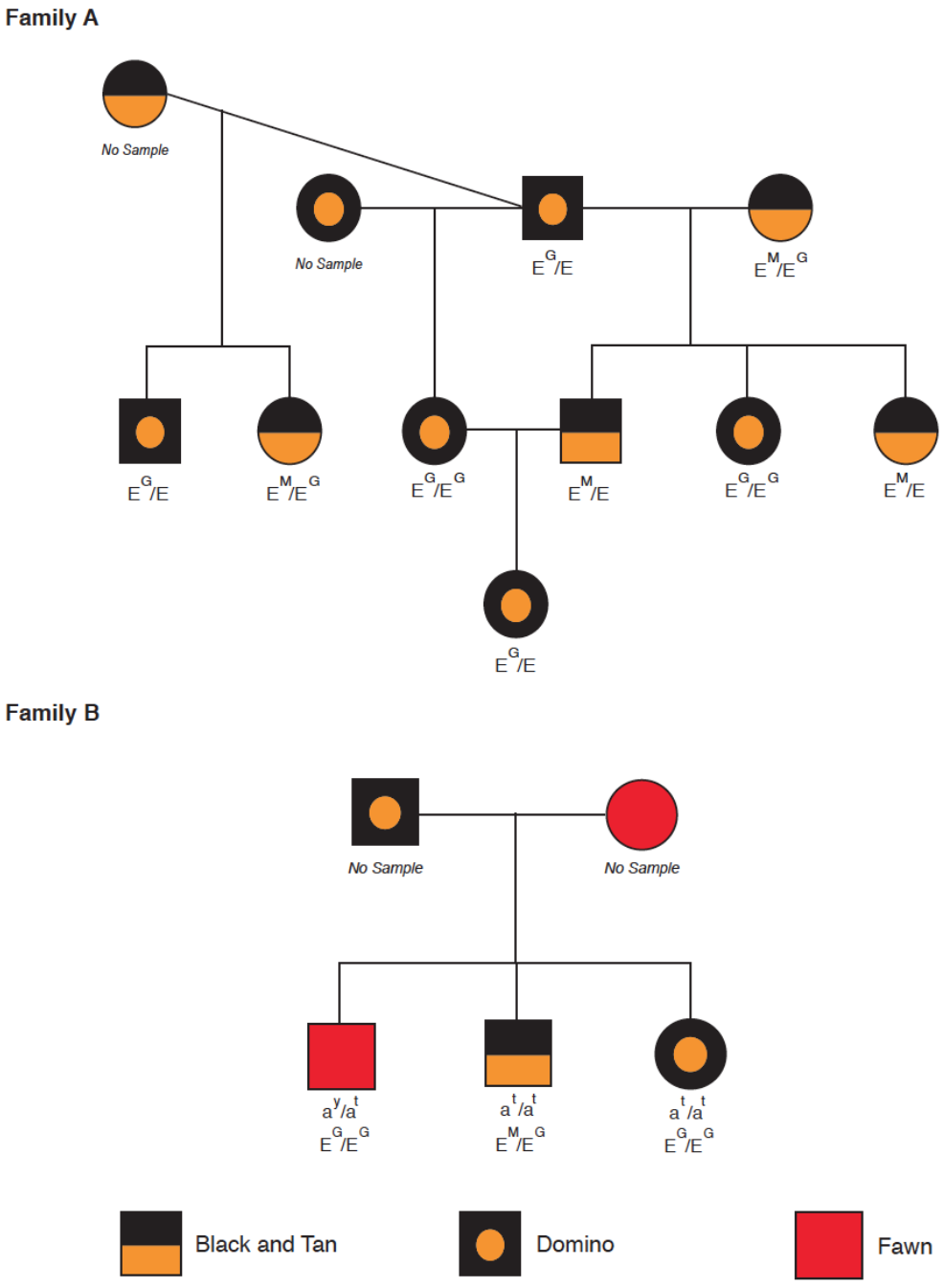


Figure 4.3: Pedigree representation of gene interactions producing the domino pattern. (A) A pedigree of an Afghan Hound family, all members of which are a^t/a^t at *ASIP* and k^y/k^y at *DEFB103*, illustrating that the E^M allele is dominant to the E^G allele. Dogs with an E^G/E^G or E^G/E genotype are domino pattern. (B) A pedigree of an Afghan Hound family that illustrates that the a^y allele at *ASIP* causes dogs to be cream to fawn, rather than domino, even with an E^G/E^G genotype.

4.5 Discussion

Little (1957) mentioned that the grizzle pattern occurred in Saluki. He suggested that it was likely a modification of sable, now known to be caused by an a^y allele. This study provides evidence that it would more appropriately be called a modification of black-and-tan. He stated that he had not studied Afghan Hounds sufficiently to postulate the alleles involved in their various colors.

Domino has been considered a rare pattern in Afghan Hounds. The pattern began to be referred to as ‘‘domino’’ because of a dog of this pattern named Tanjores Domino, born in Sweden on 10 July 1954 and imported to the United States, who did well in show circles. Prior to his popularity, the pattern was called reverse mask by fanciers of Afghan Hounds. The rarity of the domino pattern can now be better understood because this pattern requires the presence of a specific combination of alleles at more than one locus and not simply a single allele at one locus. Furthermore, the E^M allele was very common in Afghan Hounds we studied (Table 4.2). Malamutes and Siberian Huskies have a similar facial pattern and are also considered to be old breeds relatively closely related to the Afghan Hound and Saluki (Parker et al. 2004), but neither breed had the 78V allele. Some Shetland Sheepdogs and Collies have a facial pattern reminiscent of the widow’s peak appearance in Salukis and Afghan Hounds but such dogs did not have the 78V allele either.

There are dogs that are close relatives of the Saluki, known as the Sloughi and the Tazi. Dogs of these breeds may have this allele, but we did not have dogs from these breeds in our study. The 78th amino acid is highly conserved across species as a cysteine. However, in most dogs, this amino acid is a glycine and in the grizzle and domino dogs, it is a valine. Based on the postulated structure of MC1R, this amino acid is the last amino acid on the intracellular surface near the start of the second transmembrane domain (Appendix 4.1). The amino acids adjacent to the membrane appear to be important based on mutations reported in other species which affect phenotype. Based on this same 7 transmembrane structure (Garcia-Borron et al. 2005), the change from glycine to valine is postulated to dramatically alter the folding of *MCIR* mRNA, based on the 2D folding model, Quikfold (http://dinamelt.bioinfo.rpi.edu/quikfold.php). Various point mutations that cause amino acid changes, frame-shifts, and premature stop codons in *MC1R* have been reported to cause red or black coat color in several animals such as cattle (Klungland et al. 1995), horse (Marklund et al. 1996), pig (Kijas et al. 1998), dog (Newton et al.

2000), sheep (Vage et al. 2003), jaguar, and jagarundi (Eizirik et al. 2003). The E^M allele in dogs, which is caused by a g.799A>G, resulting in a valine instead of a methonine at amino acid 264 is more unusual. Eumelanin pigmented hairs are prominent on the face and sometimes also on the tail tip.

In some mammals, including most rodents, the hairs are typically composed of phaeomelanin and eumelanin bands. Some *MC1R* mutations alter the width or placement of these bands. The G78V mutation, or E^G allele, of the present study, fits better into this group of mutations. However, banded hairs occur instead of solid eumelanin pigmented hairs on the dorsal surface in grizzle Salukis and domino Afghan Hounds. McRobie et al. (2009) have reported differences in the size of phaeomelanin and eumelanin bands in the gray squirrel caused by a 24-bp deletion that includes amino acids 87–94. This mutation acts in a codominant fashion with squirrels homozygous for the deletion appearing very dark to black and squirrels heterozygous for the deletion appearing intermediate in shade. Nachman et al. (2003) reported that pale pocket mice had a subterminal band of phaeomelanin, whereas dark mice did not. This phenotypic difference is similar to the grizzle pattern in Salukis. However, the feature that is more difficult to associate with this *MC1R* mutation is the characteristic pale face with widow's peak. Given that the E^G allele only causes the grizzle or domino pattern with the widow's peak face in dogs that would otherwise be eumelanin-and-tan with an a^t/a^t genotype, we suggest that this allele minimizes the extent of the eumelanin dorsal area in such dogs. The underside color is extended up over the chest of the dog onto the face. The underside color also covers the full ventral surface, rather than the 2 chest spots and lower legs typical of most eumelanin-and-tan dogs. This mutation also appears to affect the binding of α-MSH and/or ASIP to MC1R because the dorsal surface of dogs with an a^t/a^t genotype at *ASIP* would normally consist of solid eumelanin pigmented hairs and in the presence of the E^G allele these hairs also have phaeomelanin pigmented portions.

5.0 IDENTIFICATION OF A MUTATION THAT SEGREGATES FOR THE SADDLE TAN AND BLACK-AND-TAN PHENOTYPES

5.1 Abstract

The causative mutation for the black-and-tan (a^t) phenotype in dogs was previously shown to be a SINE insertion in the 5' region of *Agouti Signaling Protein* (*ASIP*). Dogs with the black-and-tan phenotype, as well as dogs with the saddle tan phenotype, genotype as a^t /_ at this locus. We have identified a 16 bp duplication (g.1875_1890dupCCCCAGGTCAGAGTTT) in an intron of *RALY* which segregates with the black-and-tan phenotype in a group of 99 saddle tan and black-and-tan Basset Hounds and Pembroke Welsh Corgis. In these breeds, all dogs with the saddle tan phenotype had *RALY* genotypes of +/+ or +/*dup*, while dogs with the black-and-tan phenotype were homozygous for the duplication. The presence of an a^y/_ fawn or *e*/*e* red genotype is epistatic to the +/_ saddle tan genotype. Genotypes from 10 wolves and 1 coyote indicated that the saddle tan (+) allele is the ancestral allele, suggesting that black-and-tan is a modification of saddle tan. An additional 95 dogs from breeds that never have the saddle tan phenotype have all three of the possible *RALY* genotypes. I suggest that a multi-gene interaction involving *ASIP*, *RALY*, *MC1R*, *DEFB103*, and a yet-unidentified fourth gene, is required for expression of saddle tan.

5.2 Introduction

The *ASIP* gene product can act as an antagonist of α-MSH, preventing it from binding to the MC1R, resulting in the production of phaeomelanin instead of eumelanin (Lu et al. 1994). Mutations in the *ASIP* gene have been widely associated with pigmentation phenotypes in many species. Loss-of-function mutations in *MC1R* prevent ASIP from binding and produce only phaeomelanin pigment (Ollmann et al. 1998), as seen in mice (Robbins et al. 1993), cattle (Klungland et al. 1995; Joerg et al. 1996), fox (Vage et al. 1997), pigs (Kijas et al. 2001), dogs (Newton et al. 2000; Everts et al. 2000), and horses (Marklund et a. 1996). *ASIP* polymorphisms have been associated with the recessive black phenotype seen in horses (Rieder et al. 2001), sheep (Royo et al 2008), cats (Eizirik et al 2003), rats (Kuramoto et al. 2001), rabbits (Fontanesi et al. 2010), dogs (Kerns et al. 2004), and foxes (Vage et al. 1997). Eumelanin and phaeomelanin patterning have also been attributed to *ASIP* polymorphisms in brindle cattle (Girardot et al.

2006), black-and-tan pigs (Drogemuller et al. 2006), white/tan sheep (Norris et al. 2008), and black-and-tan and white-bellied agouti mice (Bultman et al. 1994; Appendix 3.1).

Four *ASIP* alleles have been identified in dogs to date, each producing a different pattern of phaeomelanin and eumelanin over the body of the dog or on the individual hair shafts (Figure 3.1). The a^y allele, caused by two adjacent amino acid substitutions of A82S R83H (Berryere et al. 2005), is the most dominant allele and produces the fawn phenotype. The wild-type allele, a^w, produces a wolf-sable phenotype. The a^t allele, caused by a SINE insertion 5' of the exon 2 start codon (Dreger and Schmutz 2011), is responsible for the black-and-tan phenotype. The *a* allele, recessive to the previous three alleles, is caused by a R96C substitution in exon 4 (Kerns et al. 2004) and results in a solid eumelanin phenotype. This allele is found primarily in herding breeds (Berryere et al. 2005).

The saddle tan phenotype is present in a limited number of dog breeds, including terriers, scent hounds, and herding dogs. The pattern involves red-based phaeomelanin pigment on the head, tail, and front and hind limbs, often reaching up to the dorsal surface at the withers and hips. Black-based eumelanin pigment is confined primarily to a saddle-shaped patch on the back. The saddle tan pattern develops as the dog ages from a puppy to a mature adult dog (Figure 5.1). As a young puppy, the color pattern strongly resembles the traditional black-and-tan phenotype (Figure 5.1), with eumelanin pigment over the majority of the body and head of the dog, and phaeomelanin pigment restricted to distinct points on the muzzle, "eyebrows", lower limbs, chest, and beneath the tail. As the dog matures, the tan areas gradually expand, reducing the area expressing eumelanin and producing the characteristic black saddle (Figure 5.1). A similar pigmentation phenotype is present in mice with the de^H allele of *TBX15*, however this color trait does not develop with age and is also accompanied by skeletal malformations (Candille et al. 2004).

Previous work (Winge 1950; Willis 1989) hypothesized that the saddle tan phenotype is caused by a separate allele, known as c^{sa} or a^s at the *Agouti Signaling Protein* (*ASIP*) locus. Recent work by our laboratory (Dreger and Schmutz 2011) indicates that dogs with the saddle tan phenotype have the same SINE insertion that is present in dogs with the black-and-tan phenotype.

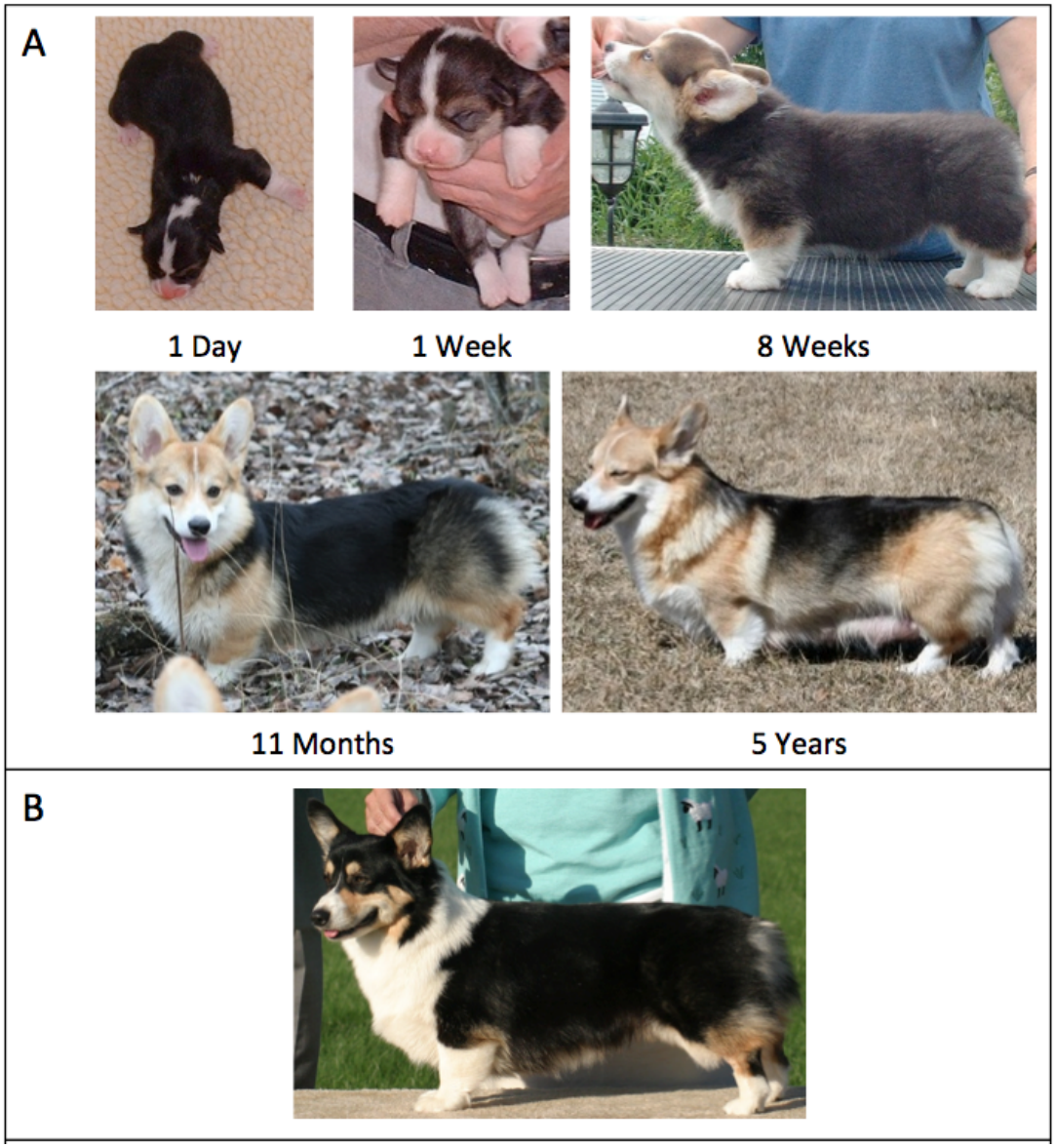


Figure 5.1: Characteristics of the saddle tan phenotype. A) The saddle tan phenotype develops as the dog ages. Nina, a saddle tan Pembroke Welsh Corgi, pictured at various ages. White spotting can obscure the tan markings on the legs, as is seen on Nina at 1 day and 1 week of age. By 8 weeks of age, the eumelanin region has receded far enough for phaeomelanin to be visible above the white markings on the legs. B) Mocha, a black-and-tan Pembroke Welsh Corgi, did not develop the saddle tan phenotype. White markings, which are common in this breed, obscure much of the phaeomelanin pigment on the front legs and chest.

During this research, a genome wide association study (GWAS) was used to map the saddle tan phenotype to a specific chromosomal region. We then used additional variants and subsequent sequencing of a smaller region to isolate a mutation that segregates with the black-and-tan phenotype in two breeds that have both the saddle tan and black-and-tan phenotypes.

5.3 Materials and Methods

5.3.1 Genome Wide Association Study

A genome wide association study was conducted using select dogs from the CanMap data set, genotyped on the Affymetrix Canine Array v2.0 SNP chip (Boyko et al. 2010) containing approximately 100K SNPs. Many dogs within the CanMap data set had owner-reported coat color information. Dogs with the black-and-tan and saddle tan patterns were selected, and verified with photographs when the reported color was ambiguous. Four breeds were considered fixed for either saddle tan (Bloodhound and Yorkshire terrier) or black and tan (Doberman Pinscher and Rottweiler). A total of 97 dogs with the black-and-tan pattern and 33 dogs with the saddle tan pattern were identified. While I designed the GWAS experiment, analysis of the GWAS data was initially conducted by Heidi Parker in the lab of Dr. Elaine Ostrander at the National Institutes of Health. Analysis was conducted using PLINK software (http://pngu.mgh.harvard.edu/~purcell/plink/), and population structure correction using EMMAX (Kang et al. 2010), on 33 saddle tan cases and 97 black-and-tan controls, and a combined 130 black-and-tan and saddle tan cases and 596 non-pattern controls.

5.3.2 Additional Dogs

DNA was obtained from Basset Hounds and Pembroke Welsh Corgis with either the black-and-tan or saddle tan phenotype using cheek brushes (Epicenter, Madison, WI). The Basset Hound and Pembroke Welsh Corgi breeds were selected because each breed displays both the saddle tan and the black-and-tan phenotypes.

The majority of the Basset Hound samples were collected from the Basset Hound Club of America national specialty show, held in Oconomowoc, WI in 2010. The remaining Basset Hound and all of the Pembroke Welsh Corgi DNA samples were obtained through contact with individual breeders and owners. DNA from our existing bank of dog DNA samples was used for validation in other breeds. The coat color and pattern of all dogs was reported by the owners and

verified with photographs. All dogs were owned by private individuals and signed consent was obtained, as is required by the Canadian Animal Care guidelines.

5.3.3 Primer Design, PCR, and Sequencing

Primers were designed to amplify regions of approximately 800 bp throughout the approximately 598 kb candidate region identified by the GWAS analysis (Appendix 3.4 b). Canine chromosome 24 shotgun sequence from the 2.0 assembly of the Boxer dog genome (Lindblad-Toh et al. 2005) was obtained from GenBank and used for designing primers.

Primers for the detection of the long and short isoforms of *RALY* mRNA were designed from predicted mRNA sequence (XM_542969) for exons 3, 5, and 9. Primers in exon 3 and 9 were used to detect the short isoform, while the long isoform was detected using primers in exons 5 and 9. mRNA was extracted from skin biopsies of 11 dogs, 1 wolf and 1 fox, dewclaws from 4 dogs, and one each of tail clipping, testis, optic nerve, iris, and retina using previously published protocols (Berryere et al. 2003) (Table 5.3).

A forward primer designed to encompass only one of the duplication sequences and a reverse primer in exon 2 of *ASIP* (Appendix 3.4 b) were designed to try to determine if mRNA product could be obtained in dogs with a *dup*/*dup* or +/+ *RALY* genotype.

A size-based PCR test was designed to detect the presence or absence of the duplication discovered in *RALY* intron 5. These primers in intron 5 were designed from sequence obtained through initial sequencing for haplotype polymorphisms.

Polymerase chain reaction (PCR) reactions totaled 15 μl, consisting of 1.5 μl 10 x PCR buffer (Fermentas), 0.3 μl of 10 mM dNTP, 0.9 μl of 25 mM $MgCl_2$, 1 μl each of the 10 pmol/μl forward and reverse primers, 0.1 μl of 5 u/μl Taq polymerase (Fermentas), 9.2 μl dH_2O, and 1 μl of roughly 50 ng/μl DNA template. PCR was conducted in Stratagene Robocycler Gradient40 machines, with 4 minutes initial denaturation at 94°C, followed by 35-37 cycles of 50 s 94°C denaturation, 50 s annealing at primer specific temperatures (Appendix 3.4 b), and 50 s 72°C extension period, followed by a final 4 min 72°C extension period. PCR products were separated on 2% agarose gels, extracted using the QIAquick gel extraction kit (Qiagen, Mississauga, ON), and sequenced at the National Research Council of Canada Plant Biotechnology Institute, using an ABI Prism 373 Sequencer (Perkin Elmer Corporation) and the Big Dye Terminator kit (Perkin

Elmer Corporation). Sequences were aligned and analyzed with the Sequencher 4.8 software program (Gene Codes Corporation, Ann Arbour, MI).

5.3.4 Genotyping

Dogs were genotyped for the a^t allele of *ASIP* using our previously described protocol (Dreger and Schmutz 2011) based on a SINE insertion in the 5' region of *ASIP*. Published protocols for genotyping for the a^y (Berryere et al. 2005), *a* (Kerns et al. 2004), E^G (Dreger and Schmutz 2011), E^M (Schmutz et al. 2003), and *e* (Schmutz et al. 2002) alleles were used for dogs suspected of expressing or carrying those alleles, based on breed. Dogs were also genotyped for the newly discovered 16 bp duplication in intron 5 of *RALY*, described herein, with the use of a size-based PCR reaction, using the methods described above.

5.4 Results

5.4.1 Inheritance

The Basset Hound samples collected included a small family of saddle tan and black-and-tan dogs. Based on the individuals in this family, it appeared that the black-and-tan pattern is recessive to saddle tan (Figure 5.2).

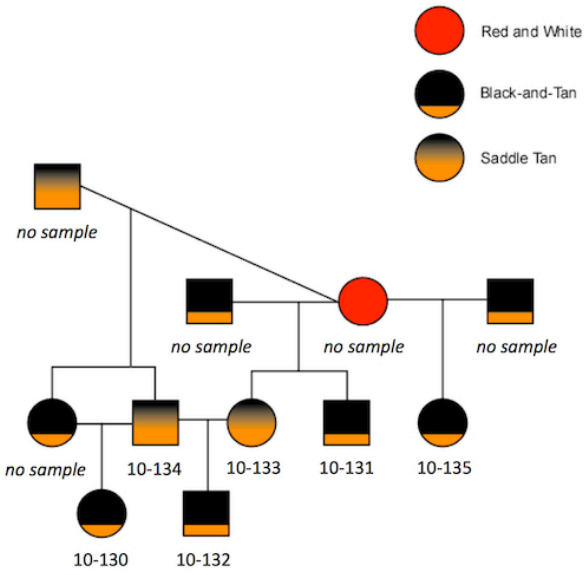


Figure 5.2: A Basset Hound family that demonstrates that a dog with the black-and-tan phenotype (10-132) can be born to two parents with the saddle tan phenotype (10-134 and 10-133).

5.4.2 GWAS Analysis

The GWAS analysis using a selection of 97 black-and-tan and 33 saddle tan dogs against 596 dogs of other phenotypes from the CanMap data set showed a significant peak consisting of two SNPs ($P_{raw} = 1.13 \times 10^{-54}$, $P_{corrected} = 2.21 \times 10^{-31}$ and $P_{raw} = 2.85 \times 10^{-52}$, $P_{corrected} = 8.96 \times 10^{-31}$) in the *ASIP* gene on chromosome 24 for all dogs that are black-and-tan or saddle tan (Figure 5.3). This confirms that the set of dogs used was sufficient to detect this coat color allele reliably and that both the black-and-tan phenotype and the saddle tan phenotype are mapped together to a genomic region that contains the SINE insertion in the *ASIP* gene, previously shown to be associated with the black-and-tan and saddle tan phenotypes (Dreger and Schmutz 2011).

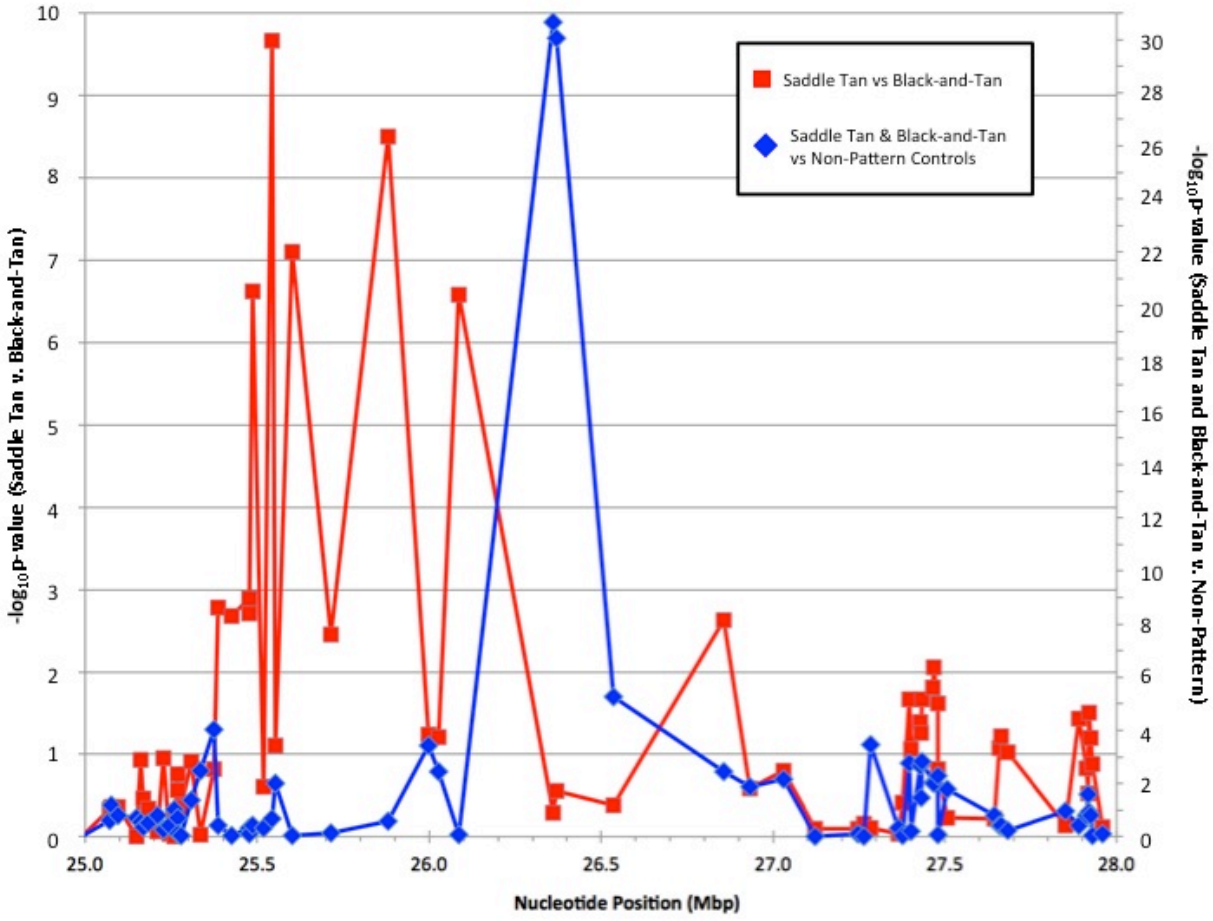


Figure 5.3: GWAS results for a narrow region on CFA24. The blue peak maps the black-and-tan and saddle tan phenotypes to a region including *ASIP*. A series of five significant red peaks map the saddle tan phenotype to a region immediately upstream of *ASIP*.

Analysis of the GWAS data from 33 saddle tan dogs against 97 dogs with the black-and-tan phenotype mapped the saddle tan phenotype to a region 5' of *ASIP* on CFA 24. A series of 5 SNPs, including nucleotides 25,486,966 ($P_{raw} = 3.9 \times 10^{-21}$, $P_{corrected} = 2.42 \times 10^{-7}$); 25,542,375 ($P_{raw} = 1.83 \times 10^{-24}$, $P_{corrected} = 2.16 \times 10^{-10}$); 25,600,559 ($P_{raw} = 3.04 \times 10^{-18}$, $P_{corrected} = 7.89 \times 10^{-8}$); 25,880,657 ($P_{raw} = 8.35 \times 10^{-22}$, $P_{corrected} = 3.15 \times 10^{-9}$); and 26,085,291 ($P_{raw} = 1.2 \times 10^{-11}$, $P_{corrected} = 2.62 \times 10^{-7}$) were selected as our primary region of interest due to the involvement of multiple significant peaks ($< P=10^{-5}$), after correction for population structure (Figure 5.3, with more detail in Appendix 5.1).

Primers designed to amplify ten segments of approximately 800 bp each throughout this region, which spans 598 kb and is approximately 870 kb upstream of the *ASIP* gene, were used in DNA sequencing assays to identify an additional ten polymorphisms. An initial group of 16 black-and-tan Basset Hounds and 18 dogs with the saddle tan pattern, consisting of 11 Basset Hounds, two each of Pembroke Welsh Corgis, Airedale Terriers, and Beagles, as well as one German Shepherd Dog, were genotyped for the five initial GWAS SNPs and the ten new polymorphisms. Genotypes were color-coded based on whether they were homozygotes (yellow or blue) or heterozygotes (green) at each locus. In this way, common haplotype blocks could be visualized and compared between the two phenotypic groups (Figure 5.4). The presence of both homozygous genotypes within a phenotype group for any given SNP indicated recombination between the SNP in question and the putative causative mutation differentiating saddle tan and black-and-tan. Accepting that the black-and-tan pattern is recessive to the saddle tan pattern, all black-and-tan dogs should have a homozygous genotype and no saddle tan dogs should have a homozygous genotype (shown in yellow in Figure 5.4). Comparison of common haplotype blocks between the two phenotype groups negated much of the initial target area, identifying a remaining candidate region including the *heterogenous nuclear ribonucleoprotein* (*hnRNP*) *associated with lethal yellow* (*RALY*) gene (Figure 5.4).

Secondary peaks in the GWAS analysis were noted on chromosomes 2 ($P = 1.14 \times 10^{-7}$), 20 ($P = 1.45 \times 10^{-5}$), and 30 ($P = 3.32 \times 10^{-6}$) (Appendix 5.2). These regions did not include candidate genes with reported function in pigment production.

Figure 5.4: Haplotype analysis for identifying a causative region for the saddle tan phenotype. Columns indicate polymorphisms, with the original SNPs from the GWAS analysis highlighted in purple in the heading. Homozygous genotypes are assigned either the blue or yellow colors and heterozygous genotypes are assigned green for purposes of visualizing genotype blocks or implied haplotypes. The symbols ‘ and “ signify multiple SNPs located within sequenced regions. Numbers beneath the region letter names indicate the nucleotide position[a] of the polymorphism.

				Region														
Lab No.	Name	Breed	Color	A	A'	A"	B	C'	C	H	H'	*STN1* ex 9	F	F'	G	D	E	*RALY* int 5	*ASIP* SNP
				25,486,966	25,487,043	25,486,073	25,542,375	25,600,520	25,600,559	25,600,942	25,601,221	25,607,959	25,640,414	25,640,464	25,711,588	25,880,657	26,085,291	26,225,244	26,359,293
10-197	George	Basset Hound	saddle tan	A/G	T/G	A/G	A/A	G/G	T/T	del/del	7/7	A/A	C/T	ins/del	C/C	G/G	C/C	+/dup	C/C
10-200	Sluggo	Basset Hound	saddle tan	A/G	T/G	A/G	A/G	A/G	T/T	ins/del	6/7	G/G	T/T	del/del	C/C	G/G	C/C	+/dup	C/C
10-218	Gabby	Basset Hound	saddle tan	A/A	T/T	G/G	G/G	A/A	T/T	ins/ins	6/6	A/A	C/C	ins/ins	C/C	G/G	C/C	+/+	C/C
10-231	Salimar	Basset Hound	saddle tan	A/A	T/T	G/G	G/G	A/G	C/T			A/G	C/T	ins/del	C/T	A/G	G/C	+/dup	C/C
10-223	Buster	Basset Hound	saddle tan	A/A	T/T	G/G	G/G	A/G	C/T	ins/ins	6/6	A/G	C/T	ins/del		A/G	G/C	+/dup	C/C
10-240	Earnest	Basset Hound	saddle tan				A/A	G/G	T/T	del/del	7/7	A/A	C/C	ins/ins	C/C	G/G	C/C	+/+	
10-192	Lena	Basset Hound	saddle tan				A/G	A/G	T/T	ins/del	6/7	A/A	C/C	ins/ins	C/C		C/C	+/+	C/C
10-216	Violet	Basset Hound	saddle tan				G/G	A/G	C/T	ins/ins	6/6	G/G	T/T	del/del	C/T	A/G	G/C	+/dup	C/C
10-209	Ginny	Basset Hound	saddle tan				G/G	A/G	C/T	ins/ins	6/6	A/G	C/T	ins/del	C/T	A/G	G/C	+/dup	C/C
10-227	Rita Red	Basset Hound	saddle tan				G/G											+/dup	
10-246	Gracie	Basset Hound	saddle tan				A/A	G/G	T/T	del/del	7/7	A/A	C/C	ins/del	C/C	G/G		+/+	C/C
10-119	Mocha	Pembroke Corgi	saddle tan				A/G	G/G	C/T	ins/del	6/7	A/G	C/T	ins/del	C/T	A/G	G/C	+/dup	C/C
10-118	Kita	Pembroke Corgi	saddle tan				A/G	G/G	C/C				C/T	ins/del	C/T	A/G	G/C	+/dup	C/C
10-262	Sasha	GSD	saddle tan				A/A	G/G	C/T			A/G	C/T	ins/del	C/T	G/G		+/+	
09-375	Chaser	Airedale Terrier	saddle tan				A/A	G/G	T/T	del/del	7/7		C/C	ins/ins	C/C	G/G	C/C	+/+	C/C
09-376	Haylee	Airedale Terrier	saddle tan				A/A			del/del	7/7	A/A	C/C	ins/ins			C/C	+/+	C/C
09-349	Sonny	Beagle	saddle tan				A/A	G/G	T/T	del/del	7/7	A/A	C/C	ins/ins	C/C	G/G	C/C	+/+	C/C
04-991	Mickey	Beagle	saddle tan				A/A	A/G	T/T	ins/del	6/7		C/T	ins/del	C/T	G/G	C/C	+/+	C/C
10-198	Bobbi	Basset Hound	black and tan	G/G	G/G	A/A	A/A									G/G	C/C		
10-241	Huey	Basset Hound	black and tan	A/G	T/G	A/G	A/G	G/G	C/T	ins/del	6/7	G/G	T/T	del/del	C/T	A/G	G/C	dup/dup	
10-239	Satin	Basset Hound	black and tan	A/G	T/G	A/G	A/G	G/G	C/T			G/G	T/T	del/del	C/T	A/G	G/C	dup/dup	C/C
10-244	Amulet	Basset Hound	black and tan	A/A	T/T	G/G	G/G	G/G	C/C	ins/ins	6/6	G/G	T/T	del/del	T/T	A/A	G/G	dup/dup	C/C
10-210	Elwood	Basset Hound	black and tan	A/G	T/G	A/G	A/G	G/G	C/T	ins/del	6/7	G/G	T/T	del/del	C/T	A/G	G/C	dup/dup	C/C
10-214	Anna	Basset Hound	black and tan	A/A	T/T	G/G	G/G	G/G	C/C	ins/ins	6/6	G/G	T/T	del/del	T/T	A/A	G/G	dup/dup	C/C
10-196	Emmy	Basset Hound	black and tan	A/A	T/T	G/G	G/G	G/G	C/C				T/T	del/del	T/T	A/A	G/G	dup/dup	
10-220	Boots	Basset Hound	black and tan	A/G	T/G	A/G	G/G	G/G	C/T	ins/del	6/7		T/T	del/del	C/T	A/G	G/C	dup/dup	C/C
10-215	Annie T	Basset Hound	black and tan				G/G	G/G	C/C	ins/ins	6/6	G/G	T/T	del/del	T/T	A/A	G/G	dup/dup	C/C
10-205	Yum Yum	Basset Hound	black and tan				A/G	G/G	C/T	ins/del	6/7	G/G	T/T	del/del	C/T		G/C	dup/dup	C/C
10-219	Giotto	Basset Hound	black and tan				G/G											dup/dup	
10-206	Pebbles	Basset Hound	black and tan				A/G	G/G	C/T	ins/del	6/7	G/G	T/T	del/del	C/T	A/G	G/C	dup/dup	C/C
10-208	Purple	Basset Hound	black and tan				G/G					G/G	T/T	del/del	T/T		G/G	dup/dup	
10-211	Penny	Basset Hound	black and tan				G/G	G/G	C/C			G/G	T/T	del/del	T/T	A/A	G/G	dup/dup	C/C
10-212	Marvin	Basset Hound	black and tan				G/G	G/G	C/C			G/G	T/T	del/del	T/T	A/A	G/G	dup/dup	C/C
10-207	Sage	Basset Hound	black and tan				G/G	G/G	C/C	ins/ins	6/6	G/G	T/T	del/del	T/T		G/G	dup/dup	C/C

[a] Position numbers are as per Canine build 2.2. For comparison, Canine build 3.1 places the *RALY* duplication at approximately nt 23,255,244.

5.4.3 Phenotypic Association

Sequencing of DNA from 34 dogs lead to the identification of a tandem duplication of 16 bp (g.1875_1890dupCCCCAGGTCAGAGTTT) in intron 5 of *RALY*. This duplication segregated with the black-and-tan phenotype in the initial set of dogs with saddle tan or black-and-tan phenotypes used for haplotype analysis (Figure 5.4). Additional Basset Hounds and Pembroke Welsh Corgis with the saddle tan (n = 64) and black-and-tan (n = 35) phenotypes, 35 dogs of four breeds fixed for the saddle tan phenotype, and 10 wolves and one coyote with the wild-type phenotype were genotyped for this duplication (Table 5.1). Nine of the 10 wolves and the single coyote were homozygous for the allele lacking the duplication, indicating that this is likely the wild-type allele (+). All dogs with the saddle tan phenotype were homozygous (+/+) for the wild-type allele, or heterozygous (+/*dup*). All Basset Hounds and Pembroke Welsh Corgis with the black-and-tan phenotype were homozygous for the duplication (*dup/dup*). Dogs from breeds that are fixed for the saddle tan phenotype, namely Airedale Terrier, Australian Terrier, Beagle, and Welsh Terrier, all had the +/+ genotype.

While most dogs with the saddle tan phenotype have little to no eumelanin on the top of their head, some dogs have moderate amounts of eumelanin on the head, occasionally forming a "widow's peak" pattern, though never extending as much as is seen with the traditional black-and-tan phenotype (Figure 5.5). In the case of Pembroke Welsh Corgis, this phenotype is referred to as "capped". Included in the original population of Pembroke Welsh Corgis and Basset Hounds genotyped for the *RALY* duplication were 5 capped saddle tan Pembroke Welsh Corgis, and 10 Basset Hounds exhibiting the same phenotype. All dogs of these breeds with the capped saddle tan phenotype genotyped as +/*dup* at *RALY*, verifying that dogs with the "capped" phenotype are more accurately grouped as a variation of the saddle tan phenotype than a variation of the black-and-tan phenotype. However, not all Pembroke Welsh Corgis or Basset Hounds with the +/*dup* genotype displayed the capped phenotype, some expressed the traditional saddle tan pattern phaeomelanin covering the majority of the head.

Table 5.1: *RALY* genotypes of dogs of breeds that segregate, or are fixed, for saddle tan or black-and-tan, and of wild canids that have neither phenotype.

Breed	**Saddle Tan Occurrence**	**Color**	***RALY* Genotypes**		
			+/+	*+/dup*	*dup/dup*
Basset Hound	Variable	Saddle Tan	13	30	0
		Black-and-Tan	0	0	27
Pembroke Welsh Corgi	Variable	Saddle Tan	6	15	0
		Black-and-Tan	0	0	8
Airedale Terrier	Fixed	Saddle Tan	15	0	0
Australian Terrier	Fixed	Saddle Tan	1	0	0
Beagle	Fixed	Saddle Tan	17	0	0
Welsh Terrier	Fixed	Saddle Tan	2	0	0
Dachshund	Not Present	Black-and-Tan	5	2	2
Doberman Pinscher	Not Present	Black-and-Tan	8	22	15
Hovawart	Not Present	Black-and-Tan	1	0	10
Shetland Sheepdog	Not Present	Black-and-Tan	5	14	11
Wolf	Not Present	Wild Type	9	1	0
Coyote	Not Present	Wild Type	1	0	0

Figure 5.5: Examples of facial patterns for red-headed saddle tan, a capped saddle tan, and a black-and-tan Pembroke Welsh Corgi. Dogs with the saddle tan phenotype have a variable amount of eumelanin present on the head, and genotype as either +/+ or +/*dup* at *RALY*. Regardless of the presence of facial eumelanin, the black-and-tan phenotype is distinct for the "capped" phenotype, and is always genotyped as *dup*/*dup* in this breed.

Ninety-five dogs of 4 breeds (Dachshund, Doberman Pinscher, Shetland Sheepdog, and Hovawart) that have the black-and-tan phenotype, but never the saddle tan phenotype, were also genotyped for the *RALY* duplication (Table 5.1). Surprisingly all three *RALY* genotypes were identified in these dogs.

With the prediction that saddle tan is produced through multiple gene interactions, in order to elucidate the potential interaction of the *RALY* duplication with other known coat color genotypes, 61 additional dogs of various breeds and coat colors were genotyped for the *RALY* duplication and, where reasonable, *ASIP*, *MC1R*, and the *K* locus (Table 5.2). Dogs with an $a^y/\_$, *a/a*, *e/e*, or $K^B/\_$ genotype did not express the saddle tan phenotype, regardless of the *RALY* genotype.

Table 5.2: Dogs of additional colors or patterns genotyped for the *RALY* duplication.

			Genotypes			
Name	**Breed**	**Color**	***RALY***	***ASIP***	***MC1R***	***DEFB103***
Twister	Basset Hound	Mahogany	+/+	a^y/a^t	E^M/e	k^y/k^y
Stachie	Basset Hound	Mahogany	+/+	a^y/a^t	E/e	k^y/k^y
BB	Basset Hound	Mahogany	+/+	a^y/a^t	E/e	k^y/k^y
Rain	Belgian Tervuren	Fawn	+/+	a^y/a^y		k^y/k^y
Maria	Belgian Tervuren	Fawn	+/+	a^y/a^y		k^y/k^y
Jake	Shetland Sheepdog	Fawn (dark)	+/+	a^y/a^y	E/E	k^y/k^y
Luka	Shetland Sheepdog	Fawn (dark)	+/*dup*	a^y/a^y	E/E	k^y/k^y
Toby	Shetland Sheepdog	Fawn (light)	+/+	a^y/a^y	E/E	k^y/k^y
Bente	Hovawart	Red	*dup*/*dup*		e/e	
Amor	Hovawart	Red	*dup*/*dup*		e/e	
Ruby	Basset Hound	Red	+/+	a^t/a^t		k^y/k^y
Colin	Keeshond	Wolf Sable (dark)	*dup*/*dup*	a^w/a^w		k^y/k^y
Kodak	Keeshond	Wolf Sable (dark)	*dup*/*dup*	a^w/a^w		k^y/k^y
Magic	Keeshond	Wolf Sable (light)	*dup*/*dup*	a^w/a^w		k^y/k^y
Bobbi	Keeshond	Wolf Sable (light)	*dup*/*dup*	a^w/a^w		k^y/k^y
Kava	Hovawart	Black	+/*dup*			$K^B/\_$
Cameo	Italian Greyhound	Blue	+/+	a^y/a^y		$K^B/\_$
Sweet One	Italian Greyhound	Blue	+/+	a^y/a^y		$K^B/\_$
Jupp	Large Munsterlander	Black and White	*dup*/*dup*	a^t/a^t		K^B/K^B
Opal	Large Munsterlander	Black and White	*dup*/*dup*			K^B/K^B
Grackle	Large Munsterlander	Black and White	*dup*/*dup*	a^t/a^t		K^B/K^B
Ibis	Large Munsterlander	Black and White	*dup*/*dup*			K^B/K^B
Logan	Eurasier	Black	+/+	a/a		
Jazz	Eurasier	Black	+/+	a/a		
Tasha	Eurasier	Black	+/+	a/a		
Shadow	Eurasier	Black	+/+	a/a		

Table 5.2: Continued

Name	Breed	Color	Genotypes			
			RALY	***ASIP***	***MC1R***	***DEFB103***
Rose	Groenendal	Black	+/+	*a*/*a*		
Rev	German Shepherd Dog	Black	+/+	*a*/*a*		
Inka	German Shepherd Dog	Black	+/+	*a*/*a*		
Carbon	German Shepherd Dog	Black	+/+	*a*/*a*		
5M	German Shepherd Dog	Black	+/+	*a*/*a*		
Tess	German Shepherd Dog	Black	+/+	*a*/*a*		
Kasa	Puli	Black	+/+	*a*/*a*		
Keddi	Puli	Black	+/+	*a*/*a*		
Raz	Puli	Black	+/+	*a*/*a*		
Toby	Schipperke	Black	+/+	*a*/*a*		
CC	Schipperke	Black	+/+	*a*/*a*		
Raven	Schipperke	Black	+/+	*a*/*a*		
Ozzy	Shetland Sheepdog	Black and White	+/+	*a*/*a*	*E*/*E*	k^y/k^y
Sketch	Shetland Sheepdog	Black and White	+/+	*a*/*a*	*E*/*E*	k^y/k^y
Painter	Shetland Sheepdog	Black and White	*dup*/*dup*	*a*/*a*	*E*/*E*	k^y/k^y
Annie	Shetland Sheepdog	Black and White	*dup*/*dup*	*a*/*a*	*E*/*E*	k^y/k^y
Polar	Samoyed	White	+/+	*a*/*a*	*e*/*e*	
Mitzi	Samoyed	White	+/+	*a*/*a*	*e*/*e*	
Trixie	Samoyed	White	+/*dup*	*a*/*a*	*e*/*e*	
Holly	Samoyed	White	+/+	*a*/*a*	*e*/*e*	
Flame	Saluki	Grizzle (dark)	*dup*/*dup*	a^t/a^t	E^G/E^G	k^y/k^y
Taz	Saluki	Grizzle (dark)	+/*dup*	a^t/a^t	E^G/E^G	k^y/k^y
Seger	Saluki	Grizzle (dark)	*dup*/*dup*	a^t/a^t	E^G/E^G	k^y/k^y
Lovey	Saluki	Grizzle (dark)	+/+	a^t/a^t	E^G/E^G	k^y/k^y
Uno	Saluki	Grizzle (light)	+/*dup*	a^t/a^t	E^G/e	k^y/k^y
Fannie	Saluki	Grizzle (light)	+/+	a^t/a^t	E^G/E^G	k^y/k^y
Spirit	Saluki	Grizzle (light)	+/*dup*	a^t/a^t	E^G/e	k^y/k^y
Eden	Saluki	Grizzle (light)	+/+	a^t/a^t	E^G/e	k^y/k^y

Previous research has identified the E^G allele of *MC1R* as responsible for the grizzle/domino phenotype in Salukis and Afghan Hounds (Dreger and Schmutz 2011). The grizzle/domino phenotype results in an expanded region of the body that displays phaeomelanin pigment in a^t/a^t black-and-tan dogs. In some ways, the grizzle/domino phenotype is similar to saddle tan, however the saddle tan phenotype generally involves expansion of the phaeomelanin pigment up to the whithers and hips of the dog and, many times, completely encompassing the head. The grizzle/domino phenotype, however, involves expansion of tan points generally to the ventral surface, and rarely extending as high as the back of the dog (Figure 4.1). Likewise, the facial tan points expand only far enough so as to produce a widow's peak pattern, with eumelanin remaining on the muzzle and top of the head. In order to determine if the saddle tan genotype and the E^G allele could interact to produce some of the variation of eumelanin expression in grizzle/domino dogs, eight Salukis with at least one copy of the E^G allele, four each of dark and pale grizzle, were also genotyped for the *RALY* duplication (Table 5.2). There did not appear to be any association between the darkness, or relative amount of eumelanin expressed, and the *RALY* genotype.

Three Shetland Sheepdogs with an a^y/a^y fawn genotype and four Keeshonds with an a^w/a^w wolf-sable genotype were designated as either "dark" or "light", depending on the overall appearance of eumelanin pigmentation on the body, and were genotyped for the *RALY* duplication. No apparent pattern of inheritance of the *RALY* alleles was detected consistent with the darkness of the fawn or wolf-sable phenotypes.

The primers used to genotype for the presence or absence of this duplication are located within intron 5, flanking the duplicated region. A product of 326 bp results from a homozygous *dup*/*dup* genotype, while the +/+ genotype results in a product of 310 bp, reflecting the duplication size of 16 bp (Appendix 5.3 a). However, heterozygous +/*dup* dogs consistently produce a gel pattern of three bands, with sizes of 326 bp, 310 bp, and a larger band that appears to be approximately 340 bp. This pattern persists when an alternate forward primer is used (Appendix 5.3 b). Searching the 3.1 dog genome assembly for the region surrounding the primers using BLAST did not yield alternate binding sites. A g.G>A1947 SNP in intron 5 is located 57 bp downstream of the duplication site and is complete linkage disequilibrium with the *RALY* g.1875_1890dupCCCCAGGTCAGAGTTT polymorphism. All 21 dog that were sequenced as +/+ for the duplication were also *G*/*G* at the SNP. All 13 dogs that were sequenced

as *dup*/*dup* for the duplication were also *A*/*A* at the SNP. Likewise, all 8 +/*dup* heterozygous dogs were *A*/*G* at the SNP.

In silico sequence analysis for the entire genomic sequence of *RALY* of the Build 3.1 version of Tasha, the Boxer who is a^y/a^y at *ASIP*, was analyzed relative to a BAC genomic sequence from a Doberman Pinscher, who is a^t/a^t at *ASIP*. No variants other than the duplication and a C>T nucleotide substitution resulting in an amino acid change from Alanine to Valine (c.Ala319Val) in exon 4 were observed. However, the T allele that is present in the Boxer sequence was not seen in any of the other mRNA sequences obtained (Table 5.3), including 18 dogs, 1 wolf, and 1 fox.

5.4.4 Characterization of RALY

The mRNA sequence obtained for *RALY* (JQ944805) from a dilute brown-and-tan Doberman Pinscher consisted of nine exons with the start codon in exon 3. This differed from the predicted *RALY* sequence (XM_542969), which does not include the observed exon 2. Similar to mouse (NM_001139513 and NM_023130) and human (NM_016732 and NM_007367) *RALY*, dogs were found to have both the long (JQ944805) and the short (JQ944804) *RALY* isoforms, that differ by the inclusion or absence of exon 5. The long isoform and short isoform, lacking exon 5, were both obtained from seven different tissues from dogs: skin, testis, dewclaw, tail, optic nerve, retina and iris (Table 5.3). Additionally, both isoforms were identified in skin of dogs with a^t/a^t, a^y/a^y, and a^w/a^w *ASIP* genotypes and skin from a fox. Therefore, no pigment phenotype had exclusively one isoform, and both isoforms are present in a variety of tissue types.

In order to determine if the duplication in *RALY* intron 5 is part of an unidentified *ASIP* non-coding exon 1, attempts to amplify a mRNA PCR product utilizing primers designed within the intron 5 duplication in *RALY* and extending to the exon 2 start of *ASIP* were performed. No product was obtained for dogs with known *dup*/*dup* or +/+ genotypes.

Table 5.3: Tissues from which mRNA for both *RALY* isoforms that differ by the inclusion of exon 5 have been identified.

Name	Breed	Color	Tissue	*ASIP* Genotype	*RALY* Genotype
Jupp	Large Munsterlander	Black and White	Skin	a^t/a^t	*dup*/*dup*
Ibis	Large Munsterlander	Black and White	Skin		*dup*/*dup*
Bobbie	Akita		Skin		+/*dup*
Sally	Beagle	Saddle Tan	Skin	a^t/a^t	+/+
Belle	Doberman Pinscher	Blue-and-Tan	Skin	a^t/a^t [a]	
Carmen	Doberman Pinscher	Brown-and-Tan	Skin	a^t/a^t [a]	
Blue	Doberman Pinscher	Albino	Skin	a^t/a^t [a]	
Libby	Doberman Pinscher	Black-and-Tan	Skin	a^t/a^t [a]	
Allie	Doberman Pinscher	Brown-and-Tan	Skin	a^t/a^t [a]	
05-173	"Sled Dog"		Skin – Dorsal		
05-173	"Sled Dog"		Skin – Ventral		
Lux	Brittany Spaniel		Skin – Dorsal		
Lux	Brittany Spaniel		Skin – Ventral		
07-038	Wolf	Wild Type	Skin	a^w/a^w	+/+
Fox	Fox		Skin		
Opal	Large Munsterlander	Black and White	Dewclaw	a^t/a^t	*dup*/*dup*
Red	Shar Pei	Blue	Dewclaw	a^y/a^y	
White	Shar Pei	Blue	Dewclaw	a^y/a^y	
Black	Shar Pei	Blue	Dewclaw	a^y/a^y	
Aussie 5	Australian Shepherd	Black-and-Tan	Tail	a^t/a^t [a]	
01-298	Mixed Breed	Brown	Testis		
04-720	Mix Breed		Optic Nerve		
04-720	Mixed Breed		Retina		
04-720	Mixed Breed		Iris		

[a] Inferred *ASIP* genotypes based on expressed coat color and breed allele frequencies.

5.5 Discussion

A number of previous GWAS studies, using similar methods and population sizes to those described here, have successfully mapped canine phenotype or disease traits and, in some instances, were followed by the identification of the causative mutations for those traits. Canine chrondrodysplasia was mapped to an expressed *FGF4* retrogene using the Affymetrix canine v2.0 SNP chip, with 95 cases from 8 breeds and 702 controls from 64 breeds (Parker et al. 2009). Using the same SNP chip, Cadieu et al. (2009) identified mutations in three genes that explain coat type variation in dogs, using the multi-breed dataset described above as well as of 96 Dachshunds of three coat varieties and 76 Portuguese Water Dogs of two coat varieties. Other GWAS studies have focused on a single breed. Using simulations based on LD patterns and allele frequencies, Lindblad-Toh et al. (2005) calculated that the probability for locus identification for a given simple Mendelian dominant trait could reach 99%, using 15,000 evenly spaced SNPs, a LOD score threshold of 5, and 100 each of case and control samples. With the same parameters, but accounting for a multigenic trait, probability of locus detection is between 50% and 97%, dependent on risk factor (Lindblad-Toh et al. 2005). These estimations may actually be conservative, as Karlsson et al. (2007) successfully mapped white spotting in dogs using a SNP array of 27,000 SNPs and approximately 10 each of case and control samples.

Our initial GWAS identified the region containing *ASIP* as significant for all black-and-tan and saddle tan phenotypes (Figure 5.3), which supports our previous finding that dogs of both black-and-tan and saddle tan phenotypes have a mutation in this gene (Dreger and Schmutz 2011). However, when the saddle tan dogs were analyzed against the black-and-tan dogs, a peak that did not precisely overlap the black-and-tan peak, but spanned a region immediately upstream of the black-and-tan peak, emerged. This identified a region separate from the coding region of *ASIP* that distinguishes the saddle tan and black-and-tan phenotypes from each other. The additional SNP analysis (Figure 5.4) identified a smaller candidate region. While the strongest candidate gene within that region was *RALY*, the initial region outlined by the five significant GWAS SNPs encompassed a number of potential positional candidate genes, including: *heterogenous nuclear ribonucleoprotein associated with lethal yellow* (*RALY*), *Antioxidant Protein 1* (*ATOX1*), *BPI fold containing family B* (*BPIFB2*), *Palate Lung and Nasal Epithelium Clone protein* (*PLUNC*), *cyclin-dependent kinase 5* (*CDK5*), *alpha-Syntrophin* (*SNTA*), *Ribosomal Protein S20* (*RPS20*), *Core-Binding Factor* (*CBFA2T2*), and *E2F transcription factor*

1 (*E2F1*). *RALY* and *SNTA* were the only two associated with pigment phenotypes in mice (http://www.informatics.jax.org/), so they were chosen as the most promising candidates to pursue. Further haplotype analysis excluded *STNA* (Figure 5.4).

Dr. Benoit Hedan of the University of Rennes (Personal Communication) indicated that a preliminary comparison of a small cohort of saddle tan versus black-and-tan dogs on the European SNP chip identified two significant SNPs on chromosome 24 (TIGRP2P314845 at nt 26,168,907 p-value = 1.861×10^{-46} and G1254F1432 at nt 26,290,526 p-value = 6.207×10^{-55}) (Appendix 5.4). These two SNPs, in linkage disequilibrium with each other, flank exons 2 to 9 of the *RALY* gene and the entire coding sequence of the *EIF2B2* gene, but do not encompass the *ASIP* gene (Appendix 5.4). This provides further support to the suggestion that the chromosomal region involving the *RALY* gene is important in modifying the saddle tan phenotype to black-and-tan.

The 16 bp duplication in intron 5 of *RALY* segregated with the black-and-tan phenotype in breeds where both the saddle tan and black-and-tan phenotypes are present, namely Basset Hounds and Pembroke Welsh Corgis. Consistent genotyping for the presence or absence of the duplication allele in 99 dogs of these two breeds has suggested that the *RALY* intron 5 duplication is sufficient to serve as a predictive tool for determining if a saddle tan Basset Hound or Pembroke Welsh Corgi carries for the black-and-tan phenotype (Table 5.1). However, a more complex mechanism for the production of the saddle tan and black-and-tan phenotypes is suggested by the occurrence of all genotypes in black-and-tan dogs of breeds where saddle tan is never present (Table 5.1).

Saddle tan was found to be dominant to black-and-tan (Figure 5.2). This dominance hierarchy was previously reported by Willis (1989), while Winge (1950) postulated that saddle tan and melanistic mask could be present simultaneously, even though he suggested both were alleles of the Agouti locus. However, melanistic mask has now been shown to be an allele of *MC1R* (Schmutz et al. 2003).

While previous researchers (Winge 1950; Willis 1989) have suggested that the saddle tan phenotype is the result of the a^s allele of *ASIP*, consideration must be equally applied to the possibility that a mutation in a separate gene interacts with the *ASIP* a^t allele to produce a modified phenotype. Recognizing that the *RALY* gene is a member of the family of heterologous nuclear ribonucleoproteins (hnRNP) and that hnRNPs function as transcription factors (Jimenez-

Garcia et al. 1993), its predicted role as a modifier of *ASIP* gene expression is appropriate. Mutations involving the *RALY* gene have been implicated in pigmentation phenotypes in both mice and quail. In mice, a 170 kb deletion that involves the entire coding region of *RALY*, leaving only the non-coding first exon and promoter region, yet the entire promoter and coding region of *ASIP*, is responsible for the A^y lethal yellow phenotype (Michaud et al. 1994). The *RALY* promoter, in the presence of the deletion allele, is proposed to cause spliced transcripts with *ASIP*, causing the yellow A^y phenotype, accompanied by obesity and insulin-resistant type II diabetes (Duhl et al. 1994; Michaud et al. 1994). Homozygous inheritance of the deletion allele results in complete negation of the *RALY* gene, leading to preimplantation lethality (Duhl et al. 1994). The fusion product resulting from the *RALY* deletion in mice (GenBank S63413) indicates that a portion of *RALY* exon 1 is spliced together with *ASIP* exon 1A and exon 2. The submitted product does not extend to the *ASIP* stop codon, but is listed as a partial sequence, so it may be assumed that a complete *ASIP* product is produced, with the added segment of *RALY* exon 1. A similar deletion, encompassing >90 kb, has been found in Japanese Quail and associated with the yellow phenotype in that species (Nadeau et al. 2008). The quail deletion involves both *RALY* and *eukaryotic translation initiation factor 2B* (*EIF2B*), one or both of which may play a role in the homozygous lethality observed with the yellow phenotype (Nadeau et al. 2008). When compared to the size and location of the 170 kb mouse deletion, it would also knock out the *EIF2B* gene, although this was not mentioned in the original research (Michaud et al. 1994). NCBI MapViewer places *EIF2B* between *RALY* and *ASIP* in the dog assembly as well, so there is a potential for its involvement in the saddle tan phenotype, though it is also possible that the knockout of *EIF2B* is responsible only for the homozygous lethal phenotype, and not related to the pigmentation phenotype seen in heterozygous animals.

The canine *RALY* intron 5 duplication is located approximately 134 kb from the only reported exon 1 of dog *ASIP* and approximately 138 kb from the dog *ASIP* start codon. The extra and inverted SINE associated with the black-and-tan phenotype (Dreger and Schmutz 2011) is located intermediate to these two points (Appendix 5.4). Regulatory elements have been shown to function over distances spanning hundreds of kilobases (Nobrega et al. 2003; Qin et al. 2004) and may even be incorporated into introns of other unrelated genes (Lettice et al. 2003), allowing for the possibility of a regulatory effect of *RALY* intron 5 including the duplication and/or SNP in *ASIP*.

The regulatory region of *ASIP* is well defined in mice, and consists of 4 alternate first exons, with exon 1A located the furthest 5' of the exon 2 start codon, a distance of 37,367 bp. Assuming that the relative distance of mouse exon 1A to exon 2 can be extrapolated to dog *ASIP*, the duplication in *RALY* intron 5 is 138,293 bp, over 100,000 bp further 5' of *ASIP* exon 2 than the mouse exon 1A. While this does not necessarily negate the potential of the *RALY* duplication to affect an *ASIP* promoter, it is well beyond the known non-coding region outlined in mouse *ASIP*. Furthermore, product was not obtained from mRNA from dogs with a *dup*/*dup* or +/+ *RALY* genotype using primers designed within the *RALY* duplication and exon 2 of *ASIP*, suggesting that the duplication region itself is not included in the coding region of *ASIP*. Together, these data suggest a stronger potential for the *RALY* duplication to act as a separate gene interaction rather than an allele of *ASIP*. The mRNA sequence of both the long and short forms of canine *RALY* were unaffected by the intron 5 duplication (Table 5.3), and therefore it is not surprising that the black-and-tan phenotype is not associated with homozygous lethality, as was the case with the A^y mouse and yellow quail phenotypes.

Regardless of whether or not the duplication in intron 5 of *RALY* affects the *RALY* gene directly or impacts expression or function of the neighboring *ASIP* gene, there is a necessity for a multiple gene interaction to produce the different saddle tan and black-and-tan phenotypes. The existence of breeds that frequently have the black-and-tan phenotype, but never the saddle tan phenotype, while expressing any of the three possible *RALY* genotypes (Table 5.1), requires that an additional modifier gene is necessary to allow for the impact of the *RALY* mutation. Accepting that the saddle tan phenotype appears to be fixed in a limited number of breeds, primarily belonging to the scent hound and terrier breed types, and occurs very infrequently in breeds not of these types, a modifier gene that has become fixed for one allele or the other among dog breeds would explain the inheritance patterns observed with the saddle tan and black-and-tan phenotypes. The ability of the genotypes of the additional modifying gene to be fixed in most breeds would suggest that the allele associated with the saddle tan phenotype occurred later in breed development, after the segregation of the hunting dog breed types, which include scent hounds and terriers, among others (Parker et al. 2004).

An *e*/*e MC1R* or a^y/_ *ASIP* genotype prevents the expression of the black-and-tan phenotype associated with the a^t *ASIP* allele. The same holds true for the saddle tan phenotype, where Basset Hounds with an *e*/*e* genotype exhibited the expected red phenotype and Basset

Hounds with an $a^y/\_$ genotype expressed the fawn phenotype, termed "mahogany" within the breed, regardless of their *RALY* genotype. As such, an *ASIP* genotype of a^t/a^t is required for both black-and-tan and saddle tan, as is at least one copy of the E^M or *E* alleles of *MC1R*. Presumably, an *ASIP* genotype of a^t/a would also allow for the expression of the saddle tan phenotype, however the only breed expected to have both the a^t and *a ASIP* alleles and the ability to express saddle tan is the German Shepherd Dog, and no appropriate dogs of that breed were included in this study. Of the 24 dogs with an *a*/*a ASIP* genotype that were genotyped for the *RALY* duplication (Table 5.2), 5 are German Shepherd Dogs, a breed that has the potential to express both the saddle tan and black-and tan phenotypes. Each of the 5 German Shepherd Dogs genotyped as +/+ saddle tan at *RALY*, but displays a normal solid eumelanin phenotype, as is expected with an *a*/*a* genotype. This further supports the need for an a^t/a^t, or possibly a^t/a, genotype at *ASIP* in order to express the saddle tan phenotype.

Since the K^B allele of the *K* locus is epistatic to the a^t black-and-tan allele of *ASIP*, so too is it epistatic to the saddle tan phenotype. The grizzle/domino phenotype associated with the E^G allele of *MC1R* (Dreger and Schmutz 2011) is similar to the saddle tan phenotype in that it allows for expansion of the tan points associated with the classic black-and-tan pattern. The grizzle/domino phenotype can be expressed with a k^{br}/k^{br} or k^{br}/k^y genotype, with brindle striping present in the expanded tan point regions. Anecdotal evidence in mixed breed dogs suggests that the saddle tan phenotype can display brindle striping in the phaeomelanin regions, suggesting that a *K* locus genotype of k^{br}/k^{br} or k^{br}/k^y is conducive to saddle tan expression. However, since none of the pure breeds used in this study commonly have the k^{br} brindle allele, this has not been substantiated. All dogs displaying the saddle tan phenotype in this research have a k^y/k^y genotype.

The identification of the 16 bp duplication in intron 5 of *RALY* that segregates with the black-and-tan phenotype in Basset Hounds and Pembroke Welsh Corgis provides for a useful diagnostic tool in these breeds. However, while the duplication has conceivably greater potential as being the causative mutation for distinguishing the black-and-tan phenotype from saddle tan, utilization of the g.G>A1947 SNP, located downstream and in complete linkage disequilibrium with the duplication, may be easier to use for diagnostic purposes.

Neither the coyote nor 9 of the 10 wolves had the RALY duplication and the tenth atypical wolf was heterozygous (+/*dup*) (Table 5.1), suggesting that the lack of the duplication is the ancestral or wild-type allele (+). It follows then that the saddle tan phenotype is ancestral to

the black-and-tan phenotype, despite the relatively limited number of breeds that currently express the saddle tan phenotype, making black-and-tan a modification of the saddle tan phenotype. The popularity of the black-and-tan phenotype across breeds and breed types is likely explained by artificial selection for the striking black-and-tan phenotype over that of saddle tan. Since the saddle tan pattern is found primarily in terriers, scent hounds, and a small number of herding breeds, this finding may also shed light on the development and relation of modern dog breeds, hinting at common ancestors between breed types.

6.0 GENERAL DISCUSSION

A large proportion of wild mammals exhibit a limited number of pigmentation phenotypes within species. Whereas, there are some species, such as the domestic dog, that display an unusually high level of variation in pigmentation type and pattern. This large degree of coat color variation is likely due, at least in part, to human selection pressure, either for functional or aesthetic purposes. In the early generations of canine domestication, selection for coat color may have been purposeful, as in the case of camouflage for hunting dogs or long range identification of herding dogs, or inadvertent, an unintentional byproduct of selection of other physical or temperament traits caused by genes in close proximity. Since the formation of purebred breed registries in the 19th century (http://www.thekennelclub.org.uk/kchistory), an increasing influence of visual aesthetics appears to be playing a role in the preferred coat color phenotypes of domestic dogs. With the increasing separation of many dog breeds from their traditional use and the growing influence of the pet dog market, the dog fancy tends to prefer flashy, unique, or striking markings in many breeds, with these preferences differing over time or among individual owners.

Many coat colors and patterns are shared among mammalian species, allowing researchers to apply knowledge across species boundaries. In the classic book by Searle (1968), common phenotypes associated with specific allele series are compared across species, in the attempt to identify prevalent or evolutionarily important genetic influences on coat color. The wild-type *A* locus phenotype, characterized by banded dorsal hairs with light or grey ventral pigment, is present in mice, rats, voles, guinea pigs, rabbits, cats, dogs, and pigs (Searle 1968). All of these species also exhibit a non-agouti phenotype of solid eumelanin pigment. A light-bellied non-agouti phenotype that is alternatively called black-and-tan, is seen in mice, rabbits, and dogs. A viable yellow phenotype is recognized in mice, dogs, and foxes (Searle 1968).

Distinction of dorsoventral pigmentation patterning, where the banding and pigment-type expression differs between the dorsal and ventral surfaces, is prevalent in many of the species discussed by Searle (1968), though is likely best understood in mice where at least two phenotypes associated with *ASIP* alleles are described as having variable dorsoventral pigment (Vrieling et al. 1994) (Figure 6.1). Vrieling et al. (1994) describe the black-and-tan (a^t) and light-bellied agouti (A^w) phenotypes as consisting of ventral hairs of solid phaeomelanin and dorsal hairs of either solid eumelanin, in the case of black-and-tan, or banded eumelanin and

Mouse *ASIP* Phenotypes	Regulatory Exons: Dorsoventral 1A/1A'	Regulatory Exons: Banding 1B/1C	Possible Dog Equivalents
 Light-Bellied Agouti (A^w)	☑	☑	
 Agouti (A)	☒	☑	
 Black-and-Tan (a^t)	☑	☒	

Figure 6.1: Mouse pigmentation phenotypes resulting from alternate expression of *ASIP* regulatory exons. Mouse A^w, light-bellied agouti, is a similar phenotype to dog a^w, wolf-sable. Mouse a^t, black-and-tan, is a similar phenotype to dog a^t, black-and-tan. (Mouse photos from http://phenome.jax.org/db/q?rtn=projects/details&syn=Jax4)

phaeomelanin, in the case of light-bellied agouti. The wild-type phenotype, the *agouti* (*A*) allele in mice, has banded dorsal and ventral hairs. This difference in pigment proportions is reportedly due to differential *ASIP* transcript expression, caused by alternate first non-coding exons (Vrieling et al. 1994) (Figure 6.1). Four alternate first exons have been identified for *ASIP* in mice, with exons 1A and 1A' associated with ventral phaeomelanin expression, and exons 1B and 1C associated with hair banding (Vrieling et al. 1994). There is a potential that two promoter regions are present in the ASIP regulatory region in mice, allowing for the production of the 1A/1A' and 1B/1C transcripts independently. Hence, as demonstrated by Vrieling et al. (1994), the 1A or 1A' transcripts can be isolated from ventral skin, but not dorsal skin, of light-bellied agouti and black-and-tan mice. The 1B or 1C transcripts can be isolated from dorsal and ventral

skin of light-bellied agouti and wild-type agouti mice, where hair banding is present, but only from ventral skin of black-and-tan mice. In the case of light-bellied agouti mice, where ventral and dorsal hairs are banded, but ventral hair is lighter in shade than dorsal hair, both the 1A/1A' and 1B/1C transcripts can be isolated from the ventral skin.

While identical mechanisms have not been identified, pigmentation patterns similar to the mouse a^t and A^w phenotypes are recognized in domestic dogs. The a^t black-and-tan phenotype in dogs is expressed as a pattern of solid eumelanin pigment on much of the torso and head of the dog, with phaeomelanin pigment restricted to distinct regions on the face, distal limbs, and perianal region, similar to the phenotype of the same name in mice (Figure 6.1). The a^w wild-type or wolf-sable phenotype consists of banded hairs on the dorsal surface and head of the dog, with the ventral regions expressing primarily phaeomelanin pigment, consistent with the phaeomelanin regions seen in dogs with the black-and-tan phenotype and similar to the A^w light-bellied agouti phenotype in mice (Figure 6.1).

To date, one non-coding first exon has been identified for *ASIP* in dogs, located 35.8 kb upstream of the exon 2 start codon (Kerns et al. 2004), and at a distance from exon 2 that is most similar to mouse exons 1B and 1C. If dog exon 1 serves a similar function as mouse exons 1B and 1C in producing a banded hair phenotype, it may be possible that a canine equivalent to mouse exons 1A and 1A', associated with dorsoventral patterning, does not contain variation. The a^w wolf-sable pattern in dogs appears to always occur in conjunction with lighter coloration on the ventral surface (Figure 6.2), similar to the light-bellied agouti (A^w) phenotype in mice. An equivalent phenotype for the mouse agouti (A) allele does not appear to exist in dogs. Taking into consideration the relative locations of the mouse 1A, 1A', 1B, and 1C exons, the SINE element associated with the dog a^t black-and-tan allele (Dreger and Schmutz, 2011) may be in close proximity to the canine equivalent of mouse exon 1B or 1C, attributed to hair banding patterns (Figure 6.3). Potential mechanisms for the production of the canine black-and-tan phenotype may revolve around the ability of the additional SINE element present in the a^t allele to interfere with transcription of the *ASIP* isoform responsible for the formation of hair banding patterns through interference or inhibition of the promoter region associated with a hair banding-specific exon, or may disrupt the sequence of a banding-specific exon itself.

Figure 6.2: The wolf-sable a^w *ASIP* phenotype consistently displays banded hairs on the dorsal surface and a pale phaeomelanin ventral surface, demonstrated here by a (A) Eurasier, (B) German Shepherd Dog, (C) Swedish Vallhund, (D) Siberian Husky, (E) Dachshund, and (F) Norwegian Elkhound.

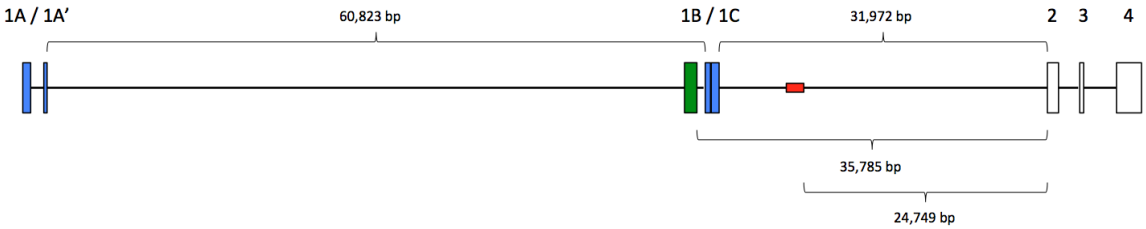


Figure 6.3: The coding and non-coding exons of *ASIP*. Clear boxes represent coding exons 2 to 4, common to mouse and dog. Blue boxes are mouse exons 1A, 1A', 1B, and 1C, as described by Vrieling et al (1994). The green box is dog exon 1, as described by Kerns et al. (2004). The red box is the SINE present in the a^t black-and-tan allele of dog.

The identification of the additional SINE element as the causative mutation for the previously postulated a^t black-and-tan allele in dogs, presented here, not only allows specific genotyping of the 4 major alleles of that gene but also expands the current understanding of the function of the *ASIP* gene in producing pigmentation patterns. This facilitates further research into gene interactions with specific *ASIP* alleles to produce distinct pigmentation phenotypes. Phenotypes characterized by eumelanin dorsal pigment and phaeomelanin ventral pigment are seen in a number of variations in domestic dogs, with each phenotype being visually distinct and identifiable, but sharing an unmistakable similarity to one another.

The grizzle phenotype in Salukis and the domino phenotype in Afghan Hounds, which are shown to be caused by the same G78V mutation in *MC1R* and denoted as the E^G allele (Dreger and Schmutz 2010), are examples of a variation on the traditional black-and-tan a^t phenotype of *ASIP*. In dogs with the grizzle/domino phenotype, the regions of phaeomelanin pigment that are characteristic of the black-and-tan phenotype expand up the legs and onto the ventral surface of the dog. The phaeomelanin cheek and eyebrow points of the black-and-tan phenotype expand and merge together, frequently resulting in a "widow's peak" of eumelanin pigment down the muzzle between the eyes (Figure 4.1). There is often an increased appearance of phaeomelanin pigment on the torso of the dog, resulting from banded or tipped hairs, rather than solid eumelanin hairs as expected with the black-and-tan phenotype.

A possible mechanism to explain the grizzle/domino phenotype is related to the physical interaction of MC1R and the ASIP ligand. The E^G mutation of *MC1R* is a substitution of amino acid 78, located on the intracellular surface of the second transmembrane domain. In the course of the grizzle/domino research, we used TMHMM *in silico* analysis (Krogh et al. 2001; Sonnhammer et al. 1998) to postulate that the E^G allele (78V) has an increased chance to alter the structure of the MC1R protein (Dreger and Schmutz 2010). This structural alteration, combined with the a^t *ASIP* allele that causes a pattern of phaeomelanin on the ventral surface, may allow for increased binding of ASIP to MC1R or decreased binding of α-MSH to MC1R, resulting in larger phaeomelanin points compared to the traditional black-and-tan phenotype. The E^M allele of *MC1R*, associated with the eumelanistic mask phenotype in dogs, is located on the extracellular surface of the sixth transmembrane domain (Schumtz et al. 2003), suggesting that mutations within the transmembrane domains of MC1R are sufficient to alter the function of the receptor.

Located three amino acids upstream of the p.Gly78Val E^G mutation is a serine residue that is part of the first intracellular loop of MC1R. An alternative prediction of function of the E^G allele might be explained by the conversion of glycine to valine at position 78 impacting the phosphorylation of that serine residue, interrupting or altering the resultant intracellular cascade that leads to the production of eumelanin pigment.

While both Salukis and Afghan Hounds have the E^G allele, producing the alternately named grizzle and domino phenotype, the perceived frequency of the phenotype appears greater in Salukis than in Afghan Hounds. This is likely due, at least in part, to the allowance of and preference for the E^M eumelanistic mask phenotype in Afghan Hounds, which is dominant to E^G domino, while the E^M allele is excluded from the breed standard for Salukis. Both Afghan Hounds and Salukis are ancient breeds (Parker et al. 2004) that originated in what is currently the Middle East as sight-hound breeds used for chasing gazelles and small game. Breed lore suggests that, traditionally, color variations within the breeds were dependent on regional climate, with dry and hot regions preferring pale colors such as cream and fawn, and cooler climates at higher elevations preferring dark colors (Goodman 2001). Modern ability to import dogs from other regions has led to a blurring of these boundaries, with a tendency for imported dogs destined for breeding in show lines to have the more visually striking color patterns (Goodman 2001). Since the grizzle/domino pattern is situated somewhere between the "light" and "dark" spectrum of possible colors for Salukis and Afghan Hounds, it is a potentially very versatile color pattern, and may have been desirable in a variety of climate regions. Dr. Adam Boyko of Cornell University has reported that the E^G allele was found in village dogs, local dogs of no particular breed type, located in Egypt, Lebanon, and Croatia (Personal Communication), suggesting that the mutation may have arisen in the Middle East or Eastern Europe. With Afghan Hounds and Salukis postulated to have been developed in the Middle East, Dr. Boyko's findings would support that the E^G allele arose in these breeds or a common ancestor of these breeds. Identification of the E^G allele in a village dog from Croatia, located in Eastern Europe, may suggest an influence of importation or transport of dogs through human migration from the Middle East to Europe. Expectation would be that, had the E^G allele arisen in Eastern European dogs, the grizzle/domino phenotype would be present in a different variety of dog breeds.

Since the publication of the Dreger and Schmutz (2010) article describing the E^G allele, a number of commercial testing laboratories have started offering a test for E^G to the public. We

have received notice that at least one Borzoi has been genotyped as E^G/E^G, indicating that this allele exists in breeds outside Salukis and Afghan Hounds. The Borzoi is a member of the sight hound family and, while it does not cluster with Afghan Hounds and Salukis in the ancient breed grouping (Parker et al. 2004), breed histories indicate that it arose from a cross between a traditional Russian breed and imported Arabian sight hounds (Fogle 2000).

Another phenotype that requires an a^t/a^t or a^t/a *ASIP* genotype, yet is visually distinct from the expected traditional black-and-tan phenotype, is saddle tan. Similar to grizzle/domino, saddle tan has phaeomelanin points that cover a greater area than is seen in dogs with the black-and-tan phenotype. However, the phaeomelanin points of the saddle tan phenotype cover an even larger area than is seen in grizzle/domino, frequently expanding up to the whithers and hip and completely covering the head of the dog (Figure 2.2). Another point that differentiates saddle tan from grizzle/domino is that it develops during the first few months of age (Figure 5.1). Pups that will eventually display the saddle tan phenotype are born with markings nearly identical to the traditional black-and-tan pattern. As the pup ages, the phaeomelanin points gradually expand in size until, at roughly one year of age, the characteristic saddle tan phenotype is expressed.

Compared again to the droopy ear phenotype in mice, caused by the *TBX15* de^H allele, the saddle tan phenotype in dogs is not accompanied by the expected skeletal malformations, decreased size, or epidermal morphology of de^H mutant mice. Rather, an intronic mutation in the *RALY* gene segregates with the age-related pigmentation phenotype in Pembroke Welsh Corgis and Basset Hounds with apparently no other impact on health or structure. *RALY*, a member of a group of *heterogenous nuclear ribonucleoprotein* (*hn-RNP*) genes, has been specifically associated with the lethal yellow phenotype in mice (Michaud et al. 1994) and quail (Nadeau et al. 2008). When quantifying hepatic levels of hn-RNP from liver tissue from rats of different ages, Brasch (1982) found that there appeared to be different protein concentrations of hn-RNPs dependent on age, with "young" rats (3-5 weeks) expressing higher levels than "mature" (8-12 weeks) and "aged" rats (over 32 weeks). The pigmentation phenotypes associated with the *RALY* gene in mice and quail are both caused by the deletion of much of the *RALY* gene, allowing the *RALY* promoter to function on the neighbouring *ASIP* gene, and leading to elevated expression of *ASIP* and the associated pigment phenotype (Michaud et al. 1994; Nadeau et al. 2008). However, the intron 5 *RALY* mutation in dogs that has been associated with the differentiation of the black-and-tan and saddle tan phenotypes does not create a deletion of the entire *RALY* gene. This,

together with the potential for age-dependent expression levels of some hn-RNPs, may indicate a direct function of canine *RALY* in creating a pigmentation phenotype that develops during maturation (Figure 5.1).

The results of my research of the genetic cause of saddle tan indicate that there is likely a complex multi-gene interaction required for the expression of the phenotype, regardless of whether the *RALY* intron 5 mutation functions directly on the *RALY* gene or as a regulatory region for *ASIP*.

As with the grizzle/domino phenotype, the potential for producing both eumelanin and phaeomelanin is required, therefore, the dog cannot have an *e*/*e* genotype at *MC1R* or a K^B allele at *DEFB103*. Likewise, there must be an a^t/a^t or a^t/a black-and-tan genotype at *ASIP*. Additionally, for the saddle tan phenotype, a dog must have a specific genotype (+/+ or +/*dup*) at *RALY* as well as a genotype at a currently unidentified modifier gene that permits the development of saddle tan. Recognizing that all three potential *RALY* genotypes appear to be evenly distributed among a selection of dog breeds that frequently exhibit the black-and-tan phenotype, but never the saddle tan phenotype, suggests that the *RALY* mutation arose prior to the mutation at the fourth gene that ultimately permitted the expression of saddle tan. This allowed the *RALY* mutation to spread among dog types undetected and without selection pressure. Further, seeing that the saddle tan pattern is almost only found in terriers and scent hounds, with a few exceptions in some herding breeds, I predict that the modifier gene acquired the mutation to permit the saddle tan phenotype after the divergence of terriers and scent hounds from other breed types. This would allow the genotypes of the modifier gene to become fixed in different populations, as suggested by the data collected herein. Parker et al. (2004) grouped terriers and scent hounds into a common ancestral grouping based on an analysis of microsatellite allele data, further supporting the hypothesis that a new mutation could become fixed in certain breed types, such as terriers and scent hounds, without influencing other unrelated breed types.

Ultimately, the differentiation of the black-and-tan, saddle tan, and grizzle/domino phenotypes is the result of modifier genes of *ASIP*. All three phenotypes require the basic ability to produce eumelanin pigment on the dorsal surface and phaeomelanin pigment on the ventral surface, but the modifier genes dictate the boundaries between the two pigment types. In the case of grizzle/domino, the G78V mutation of *MC1R* (E^G) elevates the phaeomelanin boundary to the

midpoint of the dog's body. Conversely, the wildtype *MC1R* allele (*E*) restricts the phaeomelanin expansion to smaller, more discrete points on the distal extremities and face. The saddle tan phenotype, however, appears to function in the opposite manner. The wildtype *RALY* allele (+) is associated with the saddle tan phenotype, or a greatly expanded region of phaeomelanin expression than what is considered to be the traditional black-and-tan phenotype. The *RALY* mutation, a 16 bp tandem duplication in intron 5 (*dup*), restricts the expansion of the phaeomelanin points. In this way, both *MC1R* and *RALY* can modify the phenotype caused by the a^t allele of *ASIP*, though through different directions of movement of the phaeomelanin/eumelanin boundary. The fact that a combination of an E^G grizzle and + saddle tan genotype does not create an even greater expansion of phaeomelanin pigment in the genotyped Salukis was initially surprising. However, considering that an additional modifier gene is required for the expression of saddle tan, it is possible that breeds with the E^G allele, Afghan Hounds and Salukis, do not have the correct modifier gene genotype to permit the expression of saddle tan, regardless of the *RALY* alleles present. Should a breed have both the ability to express saddle tan and the presence of the E^G allele at *MC1R*, I would predict that a phenotype would be produced that is an extreme version of the saddle tan and grizzle phenotypes. This would likely manifest as eumelanin expression only in a very small region on the back of the dog, or potentially even as a partial dorsal stripe.

The evolution of the alleles causing the various patterns resulting from *ASIP* alleles and their modification can be postulated. Using a three-gene theory, and accepting that wildtype *ASIP* is the a^w wolf-sable phenotype because only this allele is present in wolves, the grizzle/domino and saddle tan phenotypes must have arisen after domestication of the dog, as both require an a^t/a^t or a^t/a *ASIP* genotype. The a^t *ASIP* mutation, in conjunction with the wildtype forms of both *RALY* (+) and *MC1R* (*E*), cause a saddle tan phenotype. The grizzle/domino phenotype only occurs with an *MC1R* E^G allele and black-and-tan phenotype. Results of this research suggest the action of an additional modifier gene is also necessary for saddle tan. In this sense, if the wildtype form of the modifier gene allows for the black-and-tan phenotype, masking the saddle tan phenotype regardless of *RALY* genotype. Therefore, the prevailing phenotype in domestic dogs, after the introduction of the a^t *ASIP* SINE, would have been black-and-tan and not saddle tan (Table 6.1). The reverse holds true if, alternatively, the wildtype form of the modifier gene allows for the expression of the saddle tan phenotype. Once

the *ASIP* SINE is present, all dogs with the a^t *ASIP* allele and the + *RALY* allele would have the saddle tan phenotype. Mutation of the modifier gene would then mask the impact of the *RALY* mutation, resulting in increased frequency of the black-and-tan phenotype. Since the *RALY dup* allele and the *ASIP* a^t allele are both widespread across breed types, I postulate that the *RALY* and *ASIP* mutations occurred early in dog domestication. However, since the saddle tan phenotype is limited to a small number of breed types, I predict that the mutation of the modifier gene to allow for the expression of either the saddle tan or black-and-tan phenotypes, dependent on the *RALY* genotype, occurred later in breed development, at the point where terriers and scent hounds diverged from other breeds.

Table 6.1: Interaction of the genotypes of four postulated genes necessary for the production of the saddle tan phenotype.

Genotypes				Phenotype
ASIP	***MC1R***	***RALY***	**Modifier Gene**	
a^w/a^w	*E*/*E*	+/+	"*not saddle*"	Wolf Sable
a^t/a^t	*E*/*E*	+/+	"*not saddle*"	Black-and-Tan
a^t/a^t	*E*/*E*	*dup*/*dup*	"*not saddle*"	Black-and-Tan
a^t/a^t	*E*/*E*	+/+	"*saddle*"	Saddle Tan
a^t/a^t	*E*/*E*	*dup*/*dup*	"*saddle*"	Black-and-Tan

6.1 Future Research

Identifying the fourth modifier gene in this pathway is crucial for the complete understanding of the saddle tan phenotype. It may also provide valuable insight into the timeline of breed development. Specifically at what point in time did the scent hounds and terriers diverge from other breed types, and what relationships between breeds have allowed the spread of the saddle tan phenotype in a small selection of herding breeds.

I propose a genome wide association study (GWAS) as a future way of mapping the fourth modifier gene. The design of the GWAS would have to make use of dogs of specific breeds with known *RALY* and *ASIP* genotypes. Because the fourth gene appears to be fixed for one or the other homozygous genotype dependent on breed or breed type, it can be assumed that

it functions as a simple dominant/recessive, however we cannot at this point elucidate if the allele allowing for saddle tan itself is dominant or recessive. Because of this, dogs must be selected for the GWAS that have already been genotyped as a^t/a^t at *ASIP* and +/+ at *RALY* so as to limit the variability of the other contributing factors. Cohort samples must also be selected carefully from breeds with known occurrence of the black-and-tan and saddle tan phenotypes. Breeds that frequently express the black-and-tan phenotype, but never have the saddle tan phenotype, and also have all three *RALY* genotypes, can be assumed to be fixed at the allele for the fourth gene that does not allow for saddle tan. Conversely, breeds that are either fixed for saddle tan or express both the saddle tan and black-and-tan phenotypes consistent with *RALY* genotype, can be assumed to be fixed at the allele for the fourth gene that allows for saddle tan. Appropriate breeds for the saddle tan cohort would be Basset Hounds, Pembroke Welsh Corgis, and Airedale Terriers. Appropriate breeds for the black-and-tan cohort would be Shetland Sheepdogs, Dachshunds, and Doberman Pinschers. Choosing multiple breeds to represent each cohort would potentially assist in identifying a significant peak for the saddle tan gene of interest, rather than identifying genes contributing to other phenotypic variations. Since I do not predict that any one breed has both alleles at the fourth modifier gene, selecting only one breed to conduct a GWAS on would be inappropriate. Using the suggested breeds for the saddle tan and black-and-tan cohorts, would also balance out major physical characteristics such as hair type, dwarfism, overall size, and natural ear set, hopefully minimizing any potential spurious peaks. Similar to the research that was conducted for the saddle tan phenotype, any significant peaks would be screened for potential candidate genes. In the absence of strong functional candidate genes, fine mapping would be conducted through development of primers for amplification of chromosomal regions spread throughout target areas. Polymorphisms would then be identified for haplotype analysis, in hopes of refining the search area and better identifying a causative region or gene for the saddle tan modifier.

6.2 Application

A direct benefit of this research was the development of diagnostic tests for use by dog breeders, such as those offered by HealthGene (ie. http://healthgene.com/canine-dna-testing/color-testing/?tid=4&btid=25) and other labs. Diagnostic tests are now available for the a^t black-and-tan allele and the E^G grizzle/domino allele. Currently, the E^G test only applies to

Salukis, Afghan Hounds, and Borzois, where cream or fawn dogs can be tested to determine if they carry the allele for grizzle/domino.

A test for the a^t allele is of particular importance to breeds that have both the a^w wild-type and the black-and-tan phenotype. These breeds are confirmed to include German Shepherd Dogs, Border Collies, and Eurasiers, with anecdotal evidence from breeders and dog enthusiasts suggesting that breeds such as the Finnish Lapphund, "Alaskan Husky", and certain lines of Siberian Husky and Malamute may also benefit from this test.

A diagnostic test utilizing the *RALY* polymorphism would be beneficial to enthusiasts of breeds that have both the black-and-tan and saddle tan phenotypes, such as Pembroke Welsh Corgis, Basset Hounds, and potentially German Shepherd Dogs. A test result indicating the presence of the duplication allele would indicate that the dog carries the ability to pass on the black-and-tan phenotype.

6.3 Conclusion

The research compiled herein has documented a case of a three-gene interaction involving *MC1R*, *ASIP*, and *DEFB103* for the grizzle/domino phenotype, a case of a five-gene interaction involving *MC1R*, *ASIP*, *DEFB103, RALY*, and an additional modifier gene for the saddle tan phenotype. We also predicted a multiple gene interaction with the *ASIP* gene resulting in the shaded sable phenotypes.

It is becoming evident that, within a species such as the domestic dog, where pigmentation variation is so vast, one cannot hope to explain all pigmentation variation in terms of simple single-gene inheritance. The impact of multifactorial traits is widely accepted in disease research, but has often been overlooked in the seemingly simplistic scenario of pigmentation genetics, especially by dog breeders.

7.0 LITERATURE CITED

Ando H, Niki Y, Ito M, Akiyama K, Matsui MS, Yarosh DB, Ichihashi M. 2012. Melanosomes are transferred from melanocytes to keratinocytes through the process of packaging, release, uptake, and dispersion. J Invest Dermatol. 132:1222-1229.

Berryere TG, Schmutz SM, Schimpf RJ, Cowan CM, Potter J. 2003. *TYRP1* is associated with dun coat colour in Dexter cattle or how now brown cow? Anim Genet.. 34:169-175.

Berryere TG, Kerns JA, Barsh GS, Schmutz SM. 2005. Assocation of an *Agouti* allele with fawn or sable coat color in domestic dogs. Mamm Genome. 16:262-272.

Belancio VP, Hedges DJ, Deininger P. 2008. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. Genome Res. 18:343-358.

Belancio VP, Deininger PL, Roy-Engel AM. 2009. LINE dancing in the human genome: transposable elements and disease. Genome Med. 1:97-104.

Boyko AR, Boyko RH, Boyko CM, Parker HG, Castelhano M, Corey L, Degenhardt JD, Auton A, Hedimbi M, Kityo R, Ostrander EA, Schoenebeck K, Todhunter RJ, Jones P, Bustamante CD. 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. Proc Natl Acad Sci U S A. 106:13903-13908.

Boyko AR, Quignon P, Li L, Schoenebeck JL, Degenhardt JD, Lohmueller KE, Zhao K, Brisbin A, Parker HG, von Holdt BM, Cargill M, Auton A, Reynolds A, Elkahloun AG, Castelhano M, Mosher DS, Sutter NB, Johnson GS, Novembre J, Hubisz MJ, Siepel A, Wayne RK, Bustamante CD, Ostrander EA. 2010. A simple genetic architecture underlies morphological variation in dogs. PLoS Biol. 8:e1000451.

Brasch K. 1982. Age- and ploidy-related changes in rat liver nuclear proteins as assessed by one- and two-dimensional gel electrophoresis. Can J Biochem. 60: 204-214.

Bultman SJ, Michaud EJ, Woychik RP. 1992. Molecular characterization of the mouse agouti locus. Cell. 71:1195-1204.

Bultman SJ, Klebig ML, Michaud EJ, Sweet HO, Davisson MT, Woychik RP, 1994. Molecular analysis of reverse mutations from nonagouti (a) to black-and-tan (a^t) and white-bellied agouti (A^w) reveals alternative forms of agouti transcripts. Genes Dev. 8:481-90.

Cadieu E, Neff MW, Quignon P, Walsh K, Chase K, Parker HG, VonHoldt BM, Rhue A, Boyko A, Byers A, Wong A, Mosher DS, Elkahloun AG, Spady TC, Andre C, Lark KG, Cargill M, Bustamante CD, Wayne RK, Ostrander EA. 2009. Coat variation in the domestic dog is governed by variants in three genes. Science. 326:150-153.

Callinan PA, Batzer Ma. 2006. Retrotransposable elements and human disease. Genome Dyn. 1:104-115.

Candille SI, van Raamsdonk CD, Chen C, Kuijper S, Chen-Tsai Y, Russ A, Meijlink F, Barsh GS, 2004. Dorsoventral patterning of the mouse coat by Tbx15. PLoS Biol. 2:0030-0042.

Candille SI, Kaelin CB, Cattanach BM, Yu B, Thompson DA, Nix MA, Kerns JA, Schmutz SM, Millhauser GL, Barsh GS. 2007. A beta-defensin mutation causes black coat color in domestic dogs. Science. 318:1418-1423.

Chhajlani V, Wikberg JES. 1992. Molecular cloning and expression of the human melanocyte stimulating hormone receptor cDNA. FEBS Lett. 309:417-420.

Chou HH, Hayakawa T, Diaz S, Krings M, Indriati E, Leakey M, Paabo S, Satta Y, Takahata, N, Varki A, 2002. Inactivation of CMP-N-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution. Proc Natl Acad Sci U S A. 99:11736-11741.

Clark LA, Wahl JM, Rees CA, Murphy KE. 2006. Retrotransposon insertion in *SILV* is responsible for merle patterning of the domestic dog. Proc Natl Acad Sci U S A. 103:1376-1381.

Dreger DL, Schmutz SM, 2010. A new mutation in MC1R explains a coat color phenotype in 2 "old" breeds: Saluki and Afghan Hound. J Hered. 101:644-649.

Dreger DL, Schmutz SM. 2011. A SINE insertion causes the black-and-tan and saddle tan phenotypes in domestic dogs. J Hered 102:S11-S18.

Drogemuller C, Giese A, Martins-Wess F, Wiedemann S, Andersson L, Brenig B, Fries R, Leeb T. 2006. The mutation causing the black-and-tan pigmentation phenotype in Mangalitza pigs maps to the porcine ASIP locus but does not affect its coding sequence. Mamm Genome. 17:58-66.

Duhl DMJ, Stevens ME, Vrieling H, Saxon PJ, Miller MW, Epstein CJ, Barsh GS. 1994. Pleiotropic effects of the mouse lethal yellow (A^y) mutation explained by deletion of a maternally expressed gene and the simultaneous production of *agouti* fusion RNAs. Development. 120:1695-1708.

Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, O'Brien SJ. 2003. Molecular genetics and evolution of melanism in the cat family. Curr Biol. 13:448-453.

Everts RE, Rothuizen J, van Oost BA. 2000. Identification of a premature stop codon in the melanocyte-stimulating hormone receptor gene (*MC1R*) in Labrador and Golden retrievers with yellow coat color. Anim Genet. 31:194-199.

Federation Cynologique Internationale. 2010. Standards and Nomenclature. [cited 2010 December 7]. Available from: http://www.fci.be/nomenclature.aspx

Feng Q, Moran JV, Kazazian HH, Boeke JD. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell. 87:905-916.

Fogle B. 2000. The new encyclopedia of the dog. New York (NY): Dorling Kindersley.

Fontanesi L, Forestier L, Allain D, Scotti E, BerettiF, Deretz-Picoulet S, Pecchiolo E, Vernesi C, Robinson TJ, Malaney JL, Russo V, Oulmouden A. 2010. Characterization of the rabbit agouti signaling protein (*ASIP*) gene: Transcripts and phylogenetic analyses and identification of the causative mutation for the nonagouti black coat colour. Genomics 95:166-175.

Garcia-Borron J, Sanchez-Laorden BL, Jimenez-Cervantes C. 2005. Melanocotrin-1 receptor structure and functional regulation. Pigment Cell Res. 18:393-410.

Girardot M, Martin J, Guibert S, Leveziel J, Julien R, Oulmouden A, 2005. Widespread expression of the bovind Agouti gene results from at least three alternative promoters. Pigment Cell Res. 18:34-41.

Girardot M, Guibert S, Laforet MP, Gallard Y, Larroque H, Oulmouden A. 2006. The insertion of a full-length *Bos Taurus* LINE element is responsible for a transcriptional deregulation of the Normande *Agouti* gene. Pigment Cell Res. 19:346-355.

Goldberg YP, Rommens JM, Andrew SE, Hutchinson GB, Lin B, Theilmann J, Graham R, Glaves ML, Starr E, McDonald H, Nasir J, Schappert K, Kalchman MA, Clarke LA, Hayden MR. 1993. Identification of an Alu retrotransposition event in close proximity to a strong candidate gene for Huntington's disease. Nature. 362:370-373.

Goodman G. 2001. The Saluki Standard: from the Beginning. [cited 2012 May 23]. Available from www.midbarsalukis.com/standard/standard.pdf

Grall A, Guaguere E, Planchais S, Grond S, Bourrat E, Hausser In, Hitte C, Le Gallo M, Derbois C, Kim GJ, Lagoutte L, Degorce-Rubiales F, Radner FPW, Thomas A, Kury S, Bensignor E, Fontaine J, Pin D, Zimmermann R, Zechner R, Lathrop M, Galibert F, Andre C, Fischer J. 2012. PNPLA1 mutations cause autosomal recessive congenital ichthyosis in golden retriever dogs and humans. Nat Genet. 44:140-147.

Jackson IJ, Budd P, Horn JM, Johnson R, Raymond S, Steel K. 1994. Genetics and molecular biology of mouse pigmentation. Pigment Cell Res. 7:73-80.

Jimenez-Garcia LF, Spector DL. 1993. In vivo evidence that transcription and splicing are coordinated by a recruiting mechanism. Cell. 73:47-59.

Joerg H, Fries HR, Meijerink E, Stranzinger GF. 1996. Red coat color in Holstein cattle is associated with a deletion in the *MSHR* gene. Mamm Genome. 7:317-318.

Kang HM, Sul JH, Service SK, Zaitlen Na, Kong SY, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. Nat Genet. 42:348-354.

Karlsson EK, Baranowska I, Wade CM, Salmon Hillbertz NHC, Zody MC, Anderson N, Biagi TM, Patterson N, Rosengren Pielberg G, Kulbokas III EJ, Comstock KE, Keller ET, Mesirov JP, von Euler H, Kampe O, Hedhammar A, Lander ES, Andersson G, Andersson L, Lindblad-Toh K. 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. Nat Genet. 39:1321-1328.

Kerns JA, Newton J, Berryere TG, Rubin EM, Chen JF, Schmutz SM, Barsh GS. 2004. Characterization of the dog *Agouti* gene and a *nonagouti* mutation in German Shepherd Dogs. Mamm Genome. 15:798-808.

Kerns JA, Cargill EG, Clark LA, Candille SI, Berryere TG, Olivier M, Lust G, Todhunter RJ, Schmutz SM, Murphy KE, Barsh GS. 2007. Linkage and segregation analysis of black and brindle coat color in domestic dogs. Genetics. 176:1679-1689.

Kijas J, Wales R, Tornsten A, Chardon P, Moller M, Andersson L. 1998. Melanocortin Receptor 1 (*MC1R*) mutations and coat color in pigs. Genetics. 150:117-1185.

Kijas JMH, Moller M, Plastow G, Andersson L. 2001. A frameshift mutation in *MC1R* and a high frequency of somatic reversions cause black spotting in pigs. Genetics. 158:779-785.

Klungland H, Vage DI, Gomez-Raya L, Adalsteinsson S, Lien S. 1995. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. Mamm Genome. 6:636-639.

Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol. 305: 567-580.

Krull M, Brosius J, Schmitz J. 2005. Alu-SINE exonization: en route to protein-coding function. Mol Biol Evol. 22:1702-1711.

Kuramoto T, Nomoto T, Sugimura T, Ushijima T. 2001. Cloning of the rat agouti gene and identification of the rat nonagouti mutation. Mamm Genome 12:469-471

Lettice LA, Heaney SJ, Purdie LA, Li L, de Beer P, Oostra BA, Goode D, Elgar G, Hill RE, de Graaff E. 2003. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet. 12:1725-1735.

Lewin B. 2000. Genes VII. New York, NY: Oxford University Press.

Li WH, Gojobori T, Nei M, 1981. Pseudogenes as a paradigm of neutral evolution. Nature. 292: 237-239.

Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas III EJ, Zody MC, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, de Jong PJ, Kirkness E, Alvares P, Biagi T, Brockman W, Butler J, Chin CQ, Cook A, Cuff J, Daly MJ, DeCaprio D, Gnerre S, Grabherr M, Kellis M, Kleber M, Bardeleben C, Goodstadt L, Heger A, Hitte C, Kim L, Koepfli KP, Parker HG, Pollinger JP, Searle SMJ, Sutter NB, Thomas R, Webber C, Broad Institute Genome Sequencing Platform, Lander ES. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438:803-819.

Little CC. 1957. The inheritance of coat color in dogs. Ithaca, NY: Comstock Publishing Company, Inc.

Lu D, Willard D, Paten, IR, Kadwell S, Overton L, Kost T, Luther M, Chen W, Woychik RP, Wilkison WO, Cone RD. 1994. Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. Nature. 371:799-802.

Marklund L, Moller MH, Sandberg K, Andersson L. 1996. A missense mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with the chestnut coat color in horses. Mamm Genome. 7:895-899.

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis, JPA, Hirschhorn JN. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Genet. 9:356-369.

McRobie H, Thomas A, Kelly J. 2009. The genetic basis of melanism in the Gray Squirrel (*Sciurus carolinensis*). J Hered. 100:709-714.

Michaud EJ, Bultman SJ, Klebig ML, van Vugt MJ, Stubbs LJ, Russell LB, Woychik RP. 1994. A molecular model for the genetic and phenotypic characteristics of the mouse lethal yellow (A^y) mutation. Proc Natl Acad Sci U S A. 91:2562-2566.

Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr. 1996. High frequency retrotransposition in cultured mammalian cells. Cell. 87:917-927.

Mountjoy KG, Robbins LS, Mortrud MT, Cone RD. 1992. The cloning of a family of genes that encode the melanocortin receptors. Science. 257:1248-1251.

Muratani K, Kada T, Yamamoto Y, Kaneko T, Shigeto Y, Ohue T, Furuyama J, Higashino K. 1991. Inactivation of the choninesterase gene by *Alu* insertion: Possible mechanism for human gene transposition. Proc Natl Acad Sci U S A. 88:11315-11319.

Nachman MW, Hoekstra HE, D'Agosinto SL. 2003. The genetic basis of adaptive mechanism in pocket mice. Proc Natl Acad Sci U S A. 100:5268-5273.

Nadeau NJ, Minvielle F, Ito S, Inoue-Murayama M, Gourichon D, Follett SA, Burke T, Mundy NI. 2008. Characterization of Japanese Quail *yellow* as a genomic deletion upstream of the avian homolog of the mammalian *ASIP* (*agouti*) gene. Genetics. 178:777-786.

Newton J, Wilkie A, He L, Jordan S, Metallinos D, Holmes N, Jackson IJ, Barsh GS. 2000. Melanocortin 1 receptor variation in the domestic dog. Mamm Genome. 11:24-30.

Nishimura EK, Jordan SA, Oshima H, Yoshida H, Osawa M, Moriyama M, Jackson IJ, Barrandon Y, Miyachi Y, Nishikawa SI. 2002. Dominant role of the niche in melanocyte stem-cell fate determination. Nature. 416:854-859.

Nobrega MA, Ovcharenko I, Afzal V, Rubin EM. 2003. Scanning human gene deserts for long-range enhancers. Science. 302: 413.

Norris BJ, Whan VA. 2008. A gene duplication affecting expression of the ovine *ASIP* gene is responsible for white and black sheep. Genome Res. 18:1282-1293.

Ollmann MM, Lamoreux ML, Wilson BD, Barsh GS. 1998. Interaction of Agouti protein with the melanocortin 1 receptor in vitro and in vivo. Genes Dev. 12:316-330.

Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, Malek TB, Johnson GS, DeFrance HB, Ostrander EA, Kruglyak L, 2004. Genetic structure of the purebred domestic dog. Science. 304:1160-1164.

Parker HG, VonHoldt BM, Quignon P, Margulies EH, Shao S, Mosher DS, Spady TC, Elkahloun A, Cargill M, Jones PG, Maslen CL, Acland GM, Sutter NB, Kuroki K, Bustamante CD, Wayne RK, Ostrander EA. 2009. An expressed *Fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. Science 325:995-998.

Philipp U, Lupp B, Momke S, Stein V, Tipold A, Eule JC, Rehage J, Distl O. 2011. A *MITF* mutation associated with a dominant white phenotype and bilateral deafness in German Fleckvieh cattle. PLoS ONE. 6:e28857.

Prota G. 1980. Recent advances in the chemistry of melanogenesis in mammals. J Invest Dermatol. 75:122-127.

Qin Y, Kong LK, Poirier C, Truong C, Overbeek PA, Bishop CE. 2004. Long-range activation of Sox9 in Odd Sex (Ods) mice. Hum Mol Genet. 13:1213-1218.

Rieder S, Taourit S, Mariat D, Langlois B, Guerin G, 2001. Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (Equus caballus). Mamm Genome. 12:450-5.

Risch N, Merikangas K. 1996. The future of genetic studies of complex human diseases. Science. 273:1516-1517.

Ritland K, Newton C, Marshall HD. 2001. Inheritance and population structure of the white-phased "Kermode" black bear. Curr Biol. 11:1468-1472.

Robbins LS, Nadeau JH, Johnsom KR, Kelly MA, Roselli-Rehfuss L, Baack E, Mountjoy KG, Cone RD. 1993. Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. Cell. 72:827-834.

Royo LJ, Alvarez I, Arranz JJ, Fernandes I, Rodriguez AL, Perez-Pardal, Goyache F, 2008. Differences in the expression of the ASIP gene are involved in the recessive black coat colour pattern in sheep: evidence from the rare Xalda sheep breed. Anim Genet. 39:290-293.

Schmutz SM, Berryere TG, Goldfinch AD. 2002. *TYRP1* and *MC1R* genotypes and their effects on coat color in dogs. Mamm Genome. 13:380-387.

Schmutz SM, Berryere TG, Ellinwood NM, Kerns JA, Barsh GS. 2003. *MC1R* studies in dogs with melanistic mask of brindle patterns. J Hered. 94:69-73.

Schmutz SM, Berryere TG, Dreger DL. 2009. *MITF* and white spotting in dogs: A population study. J Hered. 100:S66-S74.

Schneider B, Schneider C. Eurasier vom Hirschbachtal. [cited 2010 December 6]. Available from: http://www.berndschneider100845.de

Searle AG. 1968. Comparative genetics of coat color in mammals. New York, NY: Logos Press Limited.

Seiji M, Shimao K, Birbeck MSC, Fitzpatrick TB. 1963. Subcellular localization of melanin biosynthesis. Ann N Y Acad Sci. 100:497-533.

Sinzelle L, Izsvak Z, Ivics Z. 2009. Molecular domestication of transposable elements: from detrimental parasites to useful host genes. Cell Mol Life Sci. 66:1073-1093.

Slominski A, Paus R, Plonka P, Chakraborty A, Maurer M, Pruski D, Lukiewics, S. 1994. Melanogenesis during the anagen-catagen-telogen transformation of the murine hair cycle. J Invest Dermatol. 102:862-869.

Sonnhammer ELL, von Heijne G, Krogh A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. In: Glasgow J, Littlejohn T, Major F, Lathrop R, Sankoff D, Sensen C, editors. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology; 1998, Menlo Park, CA. AAAI Press. P. 175-182.

Sorek R, Lev-Maor G, Reznik M, Dagan T, Belinky F, Gruar D, Ast G. 2004. Minimal conditions for exonization of intronic sequences: 5' splice site formation in *Alu* exons. Mol Cell. 14:221-231.

The Kennel Club. 2006 June 15. History of the Kennel Club. [cited 2012 July 18]. Available from: http://www.thekennelclub.org.uk/kchistory.

Tobin DJ, Hagen E, Botchkarev VA, Paus R. 1998. Do hair bulb melanocytes undergo apoptosis during hair follicle regression (catagen)? J Inv Dermatol. 111:941-947.

Tobin DJ. 2008. Human hair pigmentation – biological aspects. Int J Cosmet Sci. 30:233-257.

Vage DI, Lu D, Klungland H, Lien S, Adalsteinsson S, Cone RD. 1997. A non-epistatic interaction of *agouti* and *extension* in fox, *Vulpes vulpes*. Nat Genet. 15:311-315.

Vage DI, Fleet MR, Ponz R, Olsen RT, Monteagudo LV, Tejedor MT, Arruga MV, Gagilardi R, Nattrass GS, Klungland H. 2003. Mapping and characterization of the dominant black colour locus in sheep. Pigment Cell Res. 16:693-697.

Visscher PM, Brown MA, McCarthy MI, Yang J. 2012. Five years of GWAS discovery. Am J Hum Genet. 90:7-24.

Vrieling H, Duhl DMJ, Millar SE, Miller KA, Barsh GS. 1994. Differences in dorsal and ventral pigmentation result from regional expression of the mouse agouti gene. Proc Natl Acad Sci U S A. 91:5667-5671.

Wang W, Kirkness EF. 2005. Short interspersed elements (SINEs) are a major source of canine genomic diversity. Genome Res. 15:1789-1808.

Wayne RK, von Holdt BM. 2012. Evolutionary genomics of dog domestication. Mamm Genome. 23: 3-18.

Wellcome Trust Case Control Consortium. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 447:661-678.

Willis MB. 1989. Genetics of the dog. New York, NY: Howell Book House.

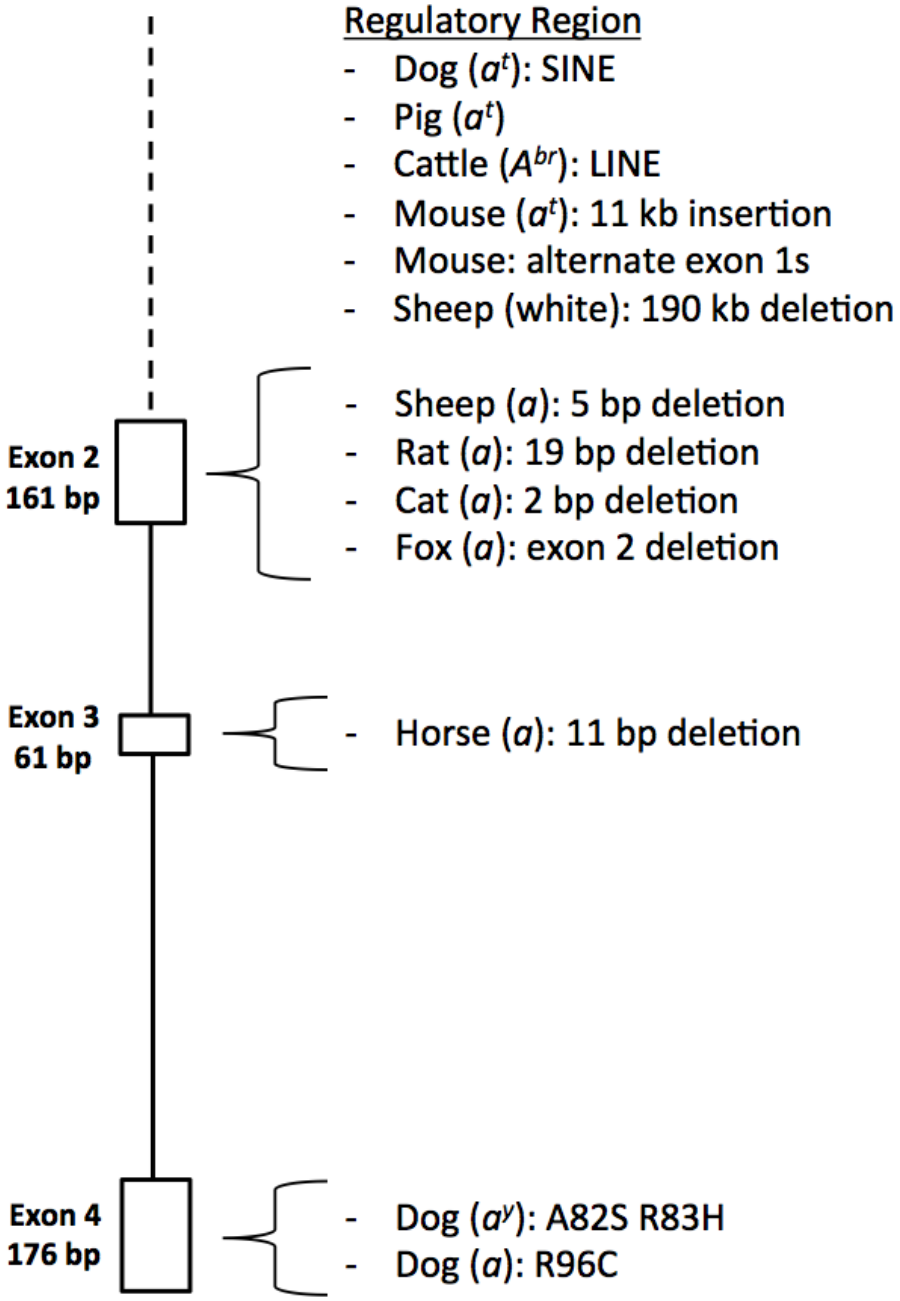
Winge O. 1950. Inheritance in dogs with special reference to hunting breeds. Ithaca, NY: Comstock Publishing Company, Inc.

Yajima I, Sato S, Kimura T, Yasumoto K, Shibahara S, Goding CR, Yamamoto H. 1999. An L1 element intronic insertion in the *black-eyed white* (*Mitf*$^{mi\text{-}bw}$) gene: the loss of a single Mitf isoform responsible for the pigmentary defect and inner ear deafness. Hum Mol Genet. 8:1431-1441.

Yasue H, Wada Y, 1996. A swine SINE (PRE-1 sequence) distribution in swine-related animal species and its phylogenetic analysis in swine genome. Anim Genet. 27:95-98.

8.0 APPENDICES

Appendix 3.1: Location of *ASIP* alleles across species



Appendix 3.2: Measurements of the length of hair pigmented with eumelanin in dogs with the a^y fawn and a^w wolf-sable phenotypes. There was no significant difference in the proportion of solid black hairs between the Keeshond and the Shetland Sheepdog breeds (Chi square = 10.64, P = 0.22)

		ASIP	% of Eumelanin per Hair										
Name	Breed	Genotype	1	2	3	4	5	6	7	8	9	10	P =
Bobbi	Keeshond	a^w/a^w	49.0	86.5	35.4	52.7	26.3	50.0	39.0	70.7	70.4	56.0	0.9988
Magic	Keeshond	a^w/a^w	49.2	42.7	100.0	46.3	100.0	100.0	32.1	88.3	41.7	39.2	0.9240
Mojo	Keeshond	a^w/a^w	57.6	65.5	82.0	62.5	60.4	54.4	100.00	63.7	56.7	80.0	0.9994
Max	Keeshond	a^w/a^w	84.6	85.6	69.2	76.1	57.7	70.0	57.9	49.1	100.0	75.2	0.9964
Genny	Keeshond	a^w/a^w	68.8	85.7	67.1	61.6	63.6	73.4	100.00	65.3	100.0	67.0	0.9977
Raye	Keeshond	a^w/a^w	86.9	70.2	100.0	82.0	52.1	72.3	81.6	68.9	71.2	75.0	0.9971
Blaze	Keeshond	a^w/a^w	100.0	35.0	51.6	100.0	40.2	100.0	100.00	100.0	45.7	100.0	0.2192
Turbo	Keeshond	a^w/a^w	57.4	47.5	100.0	100.0	100.0	100.0	100.00	57.9	100.0	52.6	0.6501
Colin	Keeshond	a^w/a^w	100.0	100.0	100.0	100.0	100.0	67.5	100.00	52.2	88.9	100.0	0.9323
Kodak	Keeshond	a^w/a^w	100.0	100.0	100.0	100.0	82.6	100.0	74.6	86.4	100.0	100.0	0.9996
Indee	Shetland Sheepdog	a^y/a^y	54.7	20.0	62.3	16.1	100.0	75.0	17.9	80.0	100.0	73.3	0.0001
Ellen	Shetland Sheepdog	a^y/a^y	59.4	74.5	60.9	47.7	57.8	69.1	47.8	55.2	63.6	69.6	0.9915
Toby	Shetland Sheepdog	a^y/a^y	100.0	36.4	34.9	70.5	36.4	51.1	81.8	50.0	34.9	100.0	0.0274
Tina	Shetland Sheepdog	a^y/a^y	100.0	43.2	100.0	100.0	47.1	44.4	100.0	36.8	100.0	45.8	0.0128
Abby*	Shetland Sheepdog	$a^y/\_$	100.0	61.1	39.4	100.0	58.9	100.0	58.3	100.0	45.1		0.2425
Bailey	Shetland Sheepdog	$a^y/\_$	58.2	100.0	46.7	60.6	76.7	51.7	77.8	81.9	100.0	85.3	0.4931
Jake	Shetland Sheepdog	a^y/a^y	100.0	76.8	41.6	100.0	85.3	79.0	100.0	100.0	100.0	100.0	0.7654
Luka	Shetland Sheepdog	a^y/a^y	100.0	100.0	100.0	86.8	100.0	50.0	100.0	100.0	100.0	100.0	1.0000
Cena	Shetland Sheepdog	a^y/a^y	37.5	36.2	39.1	42.1	39.3	43.2	40.9	37.5	36.7	40.2	1.0000

***** Only 9 hairs were analyzed for Abby due to collected sample size.

Appendix 3.3: Distribution of hair banding patterns in dogs with the a^y fawn and a^w wolf-sable phenotypes.

		Number of Hairs			
Name	**Breed**	**Solid Black**	**Banded**	**Tipped**	**Black Base**
Bobbi	Keeshond	0	8	2	0
Magic	Keeshond	3	4	3	0
Mojo	Keeshond	1	7	2	0
Max	Keeshond	1	9	0	0
Genny	Keeshond	2	8	0	0
Raye	Keeshond	1	9	0	0
Blaze	Keeshond	6	0	4	0
Turbo	Keeshond	6	1	3	0
Colin	Keeshond	7	3	0	0
Kodak	Keeshond	7	0	3	0
Indee	Shetland Sheepdog	2	5	0	3
Ellen	Shetland Sheepdog	0	8	2	0
Toby	Shetland Sheepdog	2	4	0	4
Tina	Shetland Sheepdog	5	0	0	5
Abby*	Shetland Sheepdog	4	3	0	2
Bailey	Shetland Sheepdog	2	6	0	2
Jake	Shetland Sheepdog	6	4	0	0
Luka	Shetland Sheepdog	8	2	0	0
Cena	Shetland Sheepdog	0	1	0	9

* Only 9 hairs were collected from Abby.

Appendix 3.4: Primer sequences

A) Primers used in sequencing and genotyping for the *ASIP* a^t allele mutation.

Primer Name	Sequence
#3b	5' – TTCGCCTATTCCATGGTTGC – 3'
#2	5' – ATTAACAGAGTCAAGACTAG – 3'
#5	5' – GCAGGGTAAGGTCTAGTAGT – 3'
#4	5' – TCCTCAACAAGATGATTGC – 3'
#7	5' – GCCCATCAAAAGACACCATT – 3'

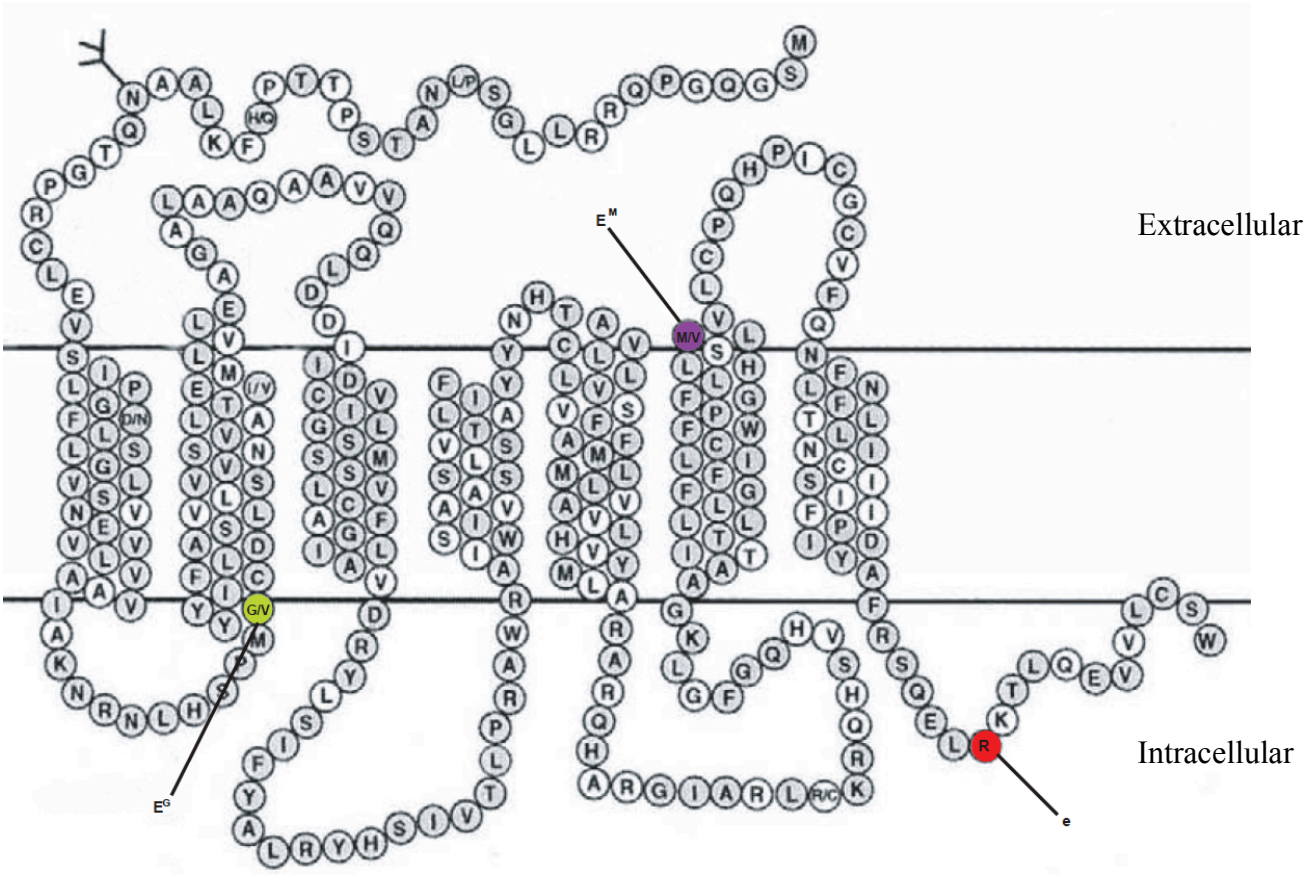
B) Primers and annealing temperatures to obtain the products that contain polymorphisms used for saddle tan haplotype analysis.

Primer Region	**Primer Sequence**	**Annealing Temperature (°C)**
Region A	F: 5'-CAGTCAGATACAGCAGCCCT-3'	59
	R: 5'-GGCTCTCATTTCACTCAGCA-3'	
Region B	F: 5'-TACATCGCCAGGCTGTCCTA-3'	57
	R: 5'-CACCTCCAGGTATGTTGAGG-3'	
Region C	F: 5'-GCTGTATACAGGGATCCTGA-3'	51
	R: 5'-TAGTCATTAGATACCTAAGG-3'	
Region D	F: 5'-GCACTCACAGCTCAAGGTCA-3'	53
	R: 5'-TCCACTTAATTTAGGAGCTC-3'	
Region E	F: 5'-GAGGCTTCATTACTTGGGCA-3'	55
	R: 5'-TCAGCGTTTCTGTCACGAAG-3'	
Region F	F: 5'-GCCTCACTCTGTTAGCTCCT-3'	51
	R: 5'-AGTATCACCTAACACATTAC-3'	
Region G	F: 5'-ATGTGTTCTTACATGCATGA-3'	53
	R: 5'-ATATCCACAGCTGTAGCTAG-3'	
Region H	F: 5'-TGGACTCAGAGTCCAGATGC-3'	51
	R: 5'-GTTAGAATCTGTGCTAGATC-3'	
Region J	F: 5'-TGAGAGATTGCTGGTGCAGC-3'	51
	R: 5'-TGGGTCAGAATCACTGAACC-3'	
Duplication Test	F: 5'-TGAGAGATTGCTGGTGCAGC-3'	51
	R: 5'-GAGACTGAACAGCAGGAGTG-'3	
mRNA Short	F: 5'-GCCACCATGTCCTTGAAGAT-3'	55
	R: 5'-CTGCTTACTGCAAGGCCCCA-3'	
mRNA Long	F: 5'-CTCTTCGACTATAGGGGCCG-3'	59
	R: 5'-CTGCTTACTGCAAGGCCCCA-3'	
RALY to *ASIP*	F: 5'-CCCCAGGTCAGAGTTT-3'	53
	R: 5'-AAGGTTCACAGAGGAGTTGC-3'	

Appendix 3.5: GenBank accession numbers resulting from this research

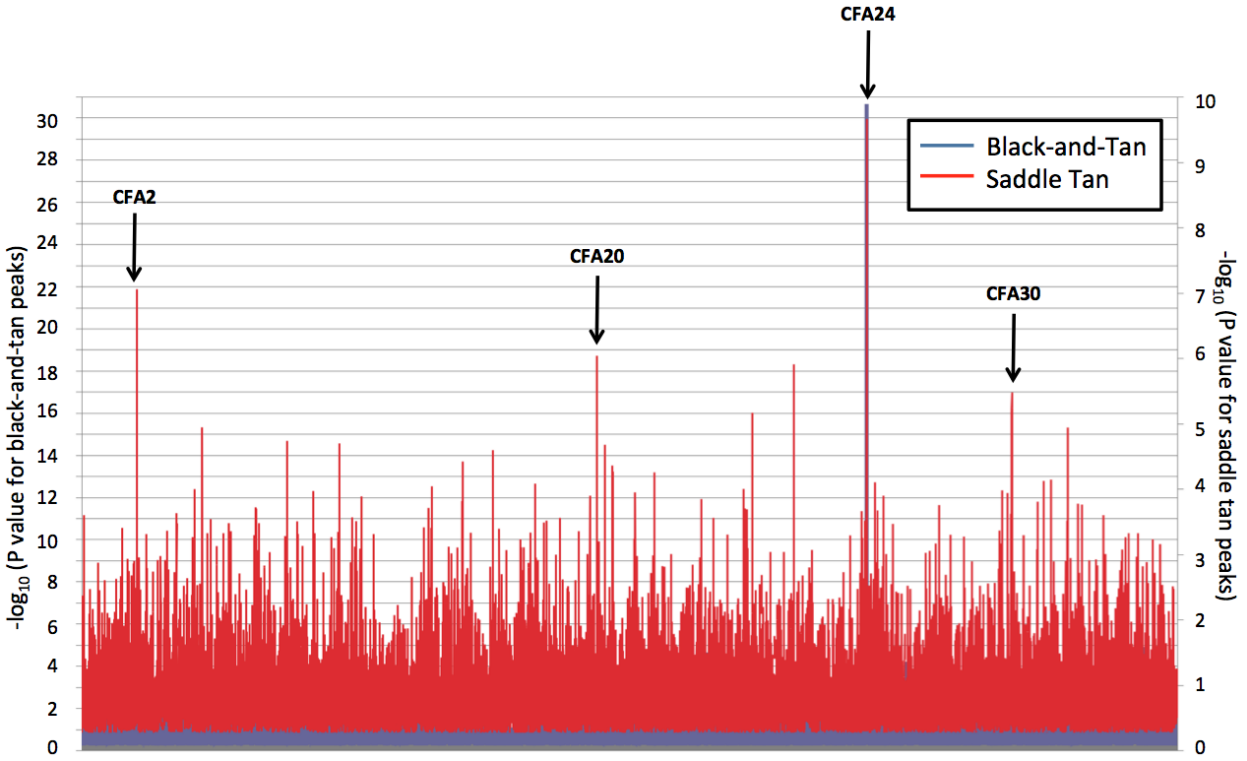
Accession Number	Description
HQ910236	Canis lupus familiaris *agouti signaling protein* (*ASIP*) gene, *ASIP a* allele, partial sequence
HQ910237	Canis lupus familiaris *agouti signaling protein* (*ASIP*) gene, *ASIP* a^t allele, partial sequence
HQ910238	Canis lupus familiaris *agouti signaling protein* (*ASIP*) gene, *ASIP* a^w allele, partial sequence
HQ910239	Canis lupus familiaris *agouti signaling protein* (*ASIP*) gene, *ASIP* a^y allele, partial sequence
GU220379	Canis lupus familiaris Afghan Hound *Melanocortin 1 Receptor* gene, complete cds
GU220378	Canis lupus familiaris Saluki *Melanocortin 1 Receptor* gene, complete cds
GU233654	Canis lupus familiaris Siberian Husky *Melanocortin 1 Receptor* gene, partial cds
GU233655	Canis lupus familiaris Siberian Husky *Melanocortin 1 Receptor* gene, complete cds
GU233656	Canis lupus familiaris Alaskan Malamute *Melanocortin 1 Receptor* gene, complete cds

Appendix 4.1: Protein structure of canine *MC1R*. Amino acids associated with the E^G, E^M, and *e* alleles are identified. (Adapted from Vage et al. 1997.)

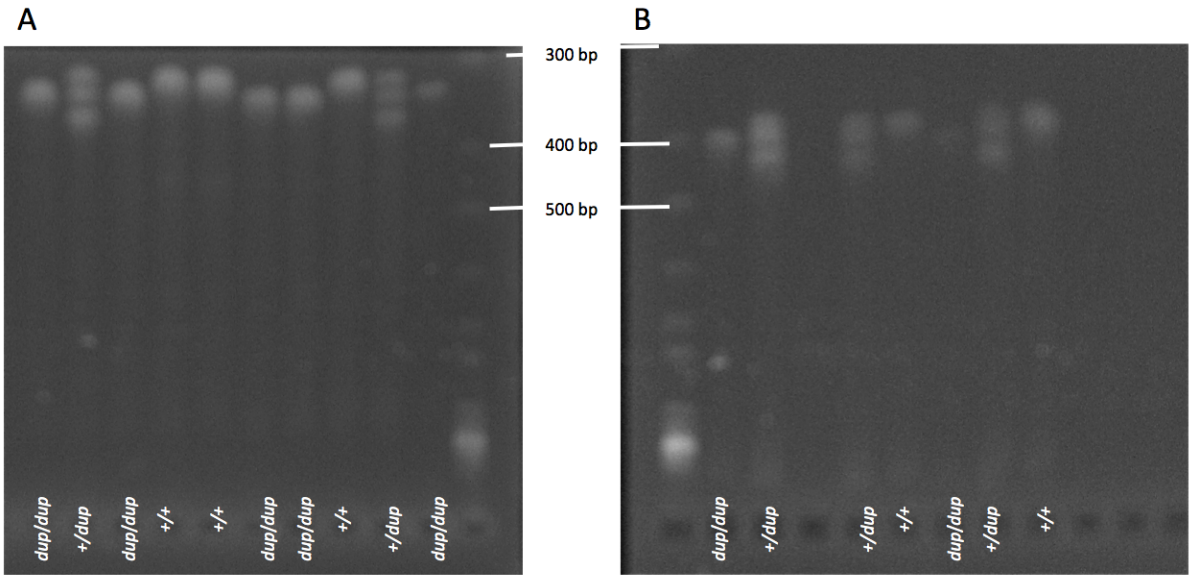


Appendix 5.1: Partial haplotype groupings for the dogs included in the initial GWAS analysis.

Appendix 5.2: Secondary GWAS peaks for the saddle tan phenotype

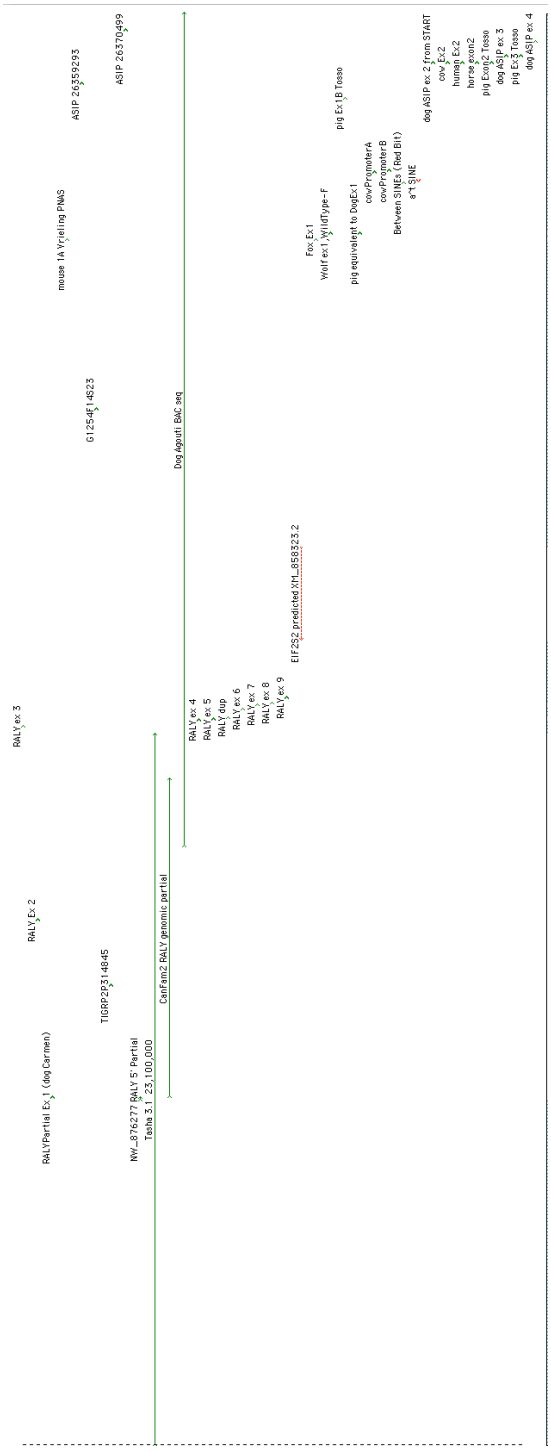


Appendix 5.3: Gel photos depicting band size and pattern for genotyping for the *RALY* duplication



Gels showing the results of the size-based test for the 16 bp duplication in *RALY* intron 5. A) Original test, with forward and reverse primers located in intron 5. B) Alternate test, with forward primer in exon 5 and the same reverse primer as in the original test. Heterozygous samples in both tests show an aberrant third band, not present in either homozygote.

Appendix 5.4: Schematic representation of the location of features relevant to *ASIP* and *RALY*[1]



[1] Points labeled TIGRP2P314845 and G1254F14S23 are the SNPS identified by Dr. Hedan as being significant for the saddle tan phenotype.